

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:07:28 ; Search time 21 Seconds
(without alignments)
487.582 Million cell updates/sec

Title: US-09-978-309A-74

Perfect score: 1195

Sequence: 1 QEKYDSMQVSLSDVTAQFES.....KKQSETKLQELNKVLGK 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/FCUTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	92.5	351	1	US-08-402-217A-2
2	1105	92.5	351	1	US-08-700-178-2
3	1105	92.5	351	3	US-08-995-654-2
4	914	76.5	606	4	US-08-477-831C-2
5	914	76.5	631	4	US-08-477-831C-11
6	913	76.4	477	1	US-08-402-217A-3
7	913	76.4	477	1	US-08-700-178-3
8	913	76.4	477	3	US-08-995-654-3
9	843	70.5	183	4	US-08-477-831C-40
10	728	60.9	183	4	US-08-477-831C-41
11	209	17.5	1388	4	US-09-572-191-2
12	209	17.5	1388	4	US-09-723-262-2
13	209	17.5	1388	4	US-09-723-219-2
14	191	16.0	976	3	US-09-104-324B-4
15	162.5	13.6	1939	4	US-09-310-187A-1
16	161.5	13.5	1886	4	US-08-938-105-3
17	159	13.3	1312	2	US-08-592-126-148
18	159	13.3	1312	2	US-08-687-080-51
19	159	13.3	1312	4	US-09-168-595-148
20	158	13.2	1388	2	US-08-685-576-4
21	156	13.1	1151	4	US-09-328-352-4744
22	153	12.8	1388	2	US-08-685-576-1
23	153	12.8	2285	4	US-09-308-375-2
24	152	12.7	2482	1	US-08-328-254-6
25	152	12.7	3248	1	US-08-353-700-1
26	152	12.7	3248	5	PCT-US95-16216-1
27	151	12.6	1354	3	US-08-685-871-2

28	149.5	12.5	2662	4	US-09-595-684B-31	Sequence 31, Appl
29	146	12.2	534	4	US-09-103-664A-2	Sequence 2, Appl
30	146	12.2	3878	4	US-09-914-259-11	Sequence 11, Appl
31	145	12.1	1196	4	US-09-107-532A-3944	Sequence 3944, Ap
32	143.5	12.0	585	4	US-09-462-951B-1	Sequence 1, Appl
33	143	12.0	1093	5	PCT-US93-03077-1	Sequence 1, Appl
34	142	11.9	1211	4	US-09-134-001C-4820	Sequence 4820, Ap
35	141	11.8	967	4	US-09-914-259-21	Sequence 21, Appl
36	140.5	11.8	288	3	US-08-312-949-4	Sequence 4, Appl
37	140.5	11.8	288	3	US-08-465-201-4	Sequence 4, Appl
38	140.5	11.8	619	1	US-08-465-746-2	Sequence 2, Appl
39	140.5	11.8	619	1	US-08-214-164-2	Sequence 2, Appl
40	140.5	11.8	619	2	US-08-467-852A-3	Sequence 3, Appl
41	140.5	11.8	619	2	US-08-245-636-2	Sequence 2, Appl
42	140.5	11.8	619	2	US-08-247-491A-3	Sequence 3, Appl
43	140.5	11.8	619	2	US-08-319-795-2	Sequence 2, Appl
44	140.5	11.8	619	3	US-08-468-985-2	Sequence 2, Appl
45	140.5	11.8	619	3	US-08-312-949-2	Sequence 2, Appl
46	140.5	11.8	648	1	US-08-072-070-2	Sequence 2, Appl
47	140.5	11.8	648	1	US-08-469-434-2	Sequence 2, Appl
48	140.5	11.8	648	1	US-08-214-222-2	Sequence 2, Appl
49	140.5	11.8	648	2	US-08-467-852A-2	Sequence 2, Appl
50	140.5	11.8	648	2	US-08-468-719-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-402-217A-2
; Sequence 2, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; TITLE OF INVENTION: HVALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,217A
; FILING DATE: 10-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF-0028US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-402-217A-2

Query Match 92.5%; Score 1105; DB 1; Length 351;
Best Local Similarity 98.7%; Pred. No. 3.2e-78;
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKYSWVOSLEDVTAQFESYKALTASIEDLKLENSLSQKAAKAGNAEDVQHOILAT 60
 DB 125 EKYSWVOSLEDVTAQFESYKALTASIEDLKLENSLSQKAAKAGNAEDVQHOILAT 184
 QY 61 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEBDFRKQLEDEGR 244
 DB 185 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEBDFRKQLEDEGR 244
 QY 121 KAEKENTTAELTEINKWRLLYEELYNKTKPFOIQDADAFEVEKQALLNEHGAQAQOLNKI 180
 DB 245 KAEKENTTAELTEINKWRLLYEELYNKTKPFOIQDADAFEVEKQALLNEHGAQAQOLNKI 304
 QY 181 RDSYAKLLGHONLKOKIKHVVKLKDENSOLKSEVSKLRCOLAKKK 225
 DB 305 RDSYAKLLGHONLKOKIKHVVKLKDENSOLKSEVSKLRCOLAKKK 349

RESULT 2

US-08-700-178-2
 ; Sequence 2, Application US/08700178
 ; Patent No. 5783669
 ; Patent No. 5783669 5700912
 ; GENERAL INFORMATION:
 ; APPLICANT: Hawkins, Phillip R.
 ; APPLICANT: Wilde, Craig G.
 ; APPLICANT: Seilhamer, Jeffrey J.
 ; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
 ; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/700,178
 ; FILING DATE: August 20, 1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/402,217
 ; FILING DATE: March 10, 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0028-1 DIV
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-845-4166
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 351 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-700-178-2

Query Match 92.5%; Score 1105; DB 1; Length 351;
 Best Local Similarity 98.7%; Pred. No. 3.2e-78;
 Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKYSWVOSLEDVTAQFESYKALTASIEDLKLENSLSQKAAKAGNAEDVQHOILAT 60
 DB 125 EKYSWVOSLEDVTAQFESYKALTASIEDLKLENSLSQKAAKAGNAEDVQHOILAT 184
 QY 61 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEBDFRKQLEDEGR 244

DB 185 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEBDFRKQLEDEGR 244
 QY 121 KAEKENTTAELTEINKWRLLYEELYNKTKPFOIQDADAFEVEKQALLNEHGAQAQOLNKI 180
 DB 245 KAEKENTTAELTEINKWRLLYEELYNKTKPFOIQDADAFEVEKQALLNEHGAQAQOLNKI 304
 QY 181 RDSYAKLLGHONLKOKIKHVVKLKDENSOLKSEVSKLRCOLAKKK 225
 DB 305 RDSYAKLLGHONLKOKIKHVVKLKDENSOLKSEVSKLRCOLAKKK 349

RESULT 3

US-08-995-654-2
 ; Sequence 2, Application US/08995654
 ; Patent No. 6025138
 ; GENERAL INFORMATION:
 ; APPLICANT: Hawkins, Phillip R.
 ; APPLICANT: Wilde, Craig
 ; APPLICANT: Seilhamer, Jeffrey
 ; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
 ; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/995,654
 ; FILING DATE: December 22, 1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/700,178
 ; FILING DATE: August 20, 1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/402,217
 ; FILING DATE: March 10, 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0028-2 DIV
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 351 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-995-654-2

Query Match 92.5%; Score 1105; DB 3; Length 351;
 Best Local Similarity 98.7%; Pred. No. 3.2e-78;
 Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QKYSWVOSLEDVTAQFESYKALTASIEDLKLENSLSQKAAKAGNAEDVQHOILAT 60
 DB 125 EKYSWVOSLEDVTAQFESYKALTASIEDLKLENSLSQKAAKAGNAEDVQHOILAT 184
 QY 61 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEBDFRKQLEDEGR 120
 DB 185 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEBDFRKQLEDEGR 244


```

US-08-402-217A-3
; Sequence 3, Application US/08402217A
; Patent No. 5587301
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,217A
; FILING DATE: 10-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF-0028US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; LIBRARY: mouse
; CLONE: GI 53979
US-08-402-217A-3

```

```

Query Match 76.4%; Score 913; DB 1; Length 477;
Best Local Similarity 76.0%; Pred. No. 3.3e-63;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;

QY 1 QEKYDSMVQSLDVTAPESYKALTASEIEDLKLENSLOEKAAGKNAEDVQHILAT 60
DB 209 QEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLENTLQEKVAAEKSVEDVQOQILTA 268
QY 61 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEDDFRQLEDEGR 120
DB 269 ESTNOEYARMVQDLQNRSLTKEEIKETITSSFLEKITDKNLQROQEDDFRQLEBKGR 328
QY 121 KAEKNTTAEITTEINKRWLLYEELYNKTKPFQIQDAPFVEKQALLNEHGAQOQLNKI 180
DB 329 TAEKENVMTEITWINKRWLLYEELYEKTKPFQOQLDAPFAEKQALLNEHGATQOQLNKI 388
QY 181 RDSYAKLLGHQNLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSETKLOEELNKVLG 240
DB 389 RDSYAQLLGHQNLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSETKLOEELNKVLG 448
QY 241 IK 242
DB 449 IR 450

```

```

RESULT 7
US-08-700-178-3
; Sequence 3, Application US/08700178

```

```

; Patent No. 5783669
; Patent No. 5783669 5700912
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,178
; FILING DATE: August 20, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,217
; FILING DATE: March 10, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0028-1 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; LIBRARY: mouse
; CLONE: GI 53979
US-08-700-178-3

```

```

Query Match 76.4%; Score 913; DB 1; Length 477;
Best Local Similarity 76.0%; Pred. No. 3.3e-63;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;

QY 1 QEKYDSMVQSLDVTAPESYKALTASEIEDLKLENSLOEKAAGKNAEDVQHILAT 60
DB 209 QEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLENTLQEKVAAEKSVEDVQOQILTA 268
QY 61 ESSNOEYVRMLDLQTKSALKETEIKETVSLQKITDLOQLKQOEDDFRQLEDEGR 120
DB 269 ESTNOEYARMVQDLQNRSLTKEEIKETITSSFLEKITDKNLQROQEDDFRQLEBKGR 328
QY 121 KAEKNTTAEITTEINKRWLLYEELYNKTKPFQIQDAPFVEKQALLNEHGAQOQLNKI 180
DB 329 TAEKENVMTEITWINKRWLLYEELYEKTKPFQOQLDAPFAEKQALLNEHGATQOQLNKI 388
QY 181 RDSYAKLLGHQNLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSETKLOEELNKVLG 240
DB 389 RDSYAQLLGHQNLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSETKLOEELNKVLG 448
QY 241 IK 242
DB 449 IR 450

```

```

RESULT 8

```



```

US-08-995-654-3
; Sequence 3, Application US/08995654
; Patent No. 6025138
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Wilde, Craig
; APPLICANT: Seilhamer, Jeffrey
; TITLE OF INVENTION: HYALURONAN RECEPTOR EXPRESSED IN HUMAN
; TITLE OF INVENTION: UMBILICAL VEIN ENDOTHELIAL CELLS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,654
; FILING DATE: December 22, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/700,178
; FILING DATE: August 20, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/402,217
; FILING DATE: March 10, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0028-2 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-995-654-3
Query Match 76.4%; Score 913; DB 3; Length 477;
Best Local Similarity 76.0%; Pred. No. 3.3e-63;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;

Qy 1 QEKYDSMVQSLDVTAFESYKALTASIEDLKLNSLQEKAAKAGKNAEDVQHQILAT 60
Db 209 QEKYNDTAQSLRDVSAQLESYKSTLKEIEDLKLNTLQEKVMAEKSVEDVQOQILTA 268

Qy 61 ESSNQEVVRMLDLQTSALKETEIKETVSLQKITDQLNQLKQOEEDFRKQLEDEGR 120
Db 269 ESTNQEVARMVQDLQNRSTLKEEIKETTSSTFLEKITDKNLQKQOEEDFRKQLEEKGR 328

Qy 121 KAEKENTTAELTEINKWRLLYEHLNKTKEFQOLDFAFEVKQALLNEHGAQEQLNKI 180
Db 329 TAEKENVMTETLMEINKWRLLYEHLNKTKEFQOLDFAFEVKQALLNEHGAQEQLNKI 388

Qy 181 RDSYAKLLGHONLKQIKHVKVLDKNSQLKSEVSKLRCLAKKQSTKLQELNKLVLG 240
Db 389 RDSYAQLLGHONLKQIKHVKVLDKNSQLKSEVSKLRSLVKKRQNELRLQGGELDKALG 448

Qy 241 IK 242
Db 449 IR 450

```

```

RESULT 9
US-08-477-831C-40
; Sequence 40, Application US/08477831C
; Patent No. 6429291
; GENERAL INFORMATION:
; APPLICANT: TURLEY, EVA A.
; APPLICANT: SHUWEN, ZHANG
; APPLICANT: ENTWISTLE, JOYCELYN
; TITLE OF INVENTION: HYALURONAN RECEPTOR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10020-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Rel. #1.0, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,831C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PIERRI, MARGARET A.
; REGISTRATION NUMBER: 30,709
; REFERENCE/DOCKET NUMBER: SIM-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 183 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: /desc = protein fragment
;
US-08-477-831C-40
Query Match 70.5%; Score 843; DB 4; Length 183;
Best Local Similarity 92.9%; Pred. No. 2.8e-58;
Matches 170; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 29 IEDLKLENSLQEKAAKAGKNAEDVQHQILATSSNOEYVRMLDLQTSALKETEIKI 88
Db 1 IEDLKLENSLQEKAAKAGKNAEDVQHQILATSSNOEYVRMLDLQTSALKETEIKI 60

Qy 89 TVSFLOKITDQLNQLKQOEEDFRKQLEDEGRKAKEKENTTAELTEINKWRLLYEELYNK 148
Db 61 TVSFLOKITDQLNQLKQOEEDFRKQLEDEGRKAKEKENTPELTMEINKWRLLYDELYEK 120

Qy 149 TKFQFOLDFAFEVKQALLNEHGAQEQLNKIRDSYAKLLGHONLKQIKHVKVLDKNS 208
Db 121 TKFPQQLDAFEVKQALLDEHGAQEQLNKIRDSYAKLLGHONLKQIKHVKVLDKNS 180

Qy 209 QLK 211
Db 181 QLK 183

RESULT 10
US-08-477-831C-41
; Sequence 41, Application US/08477831C
; Patent No. 6429291
; GENERAL INFORMATION:
; APPLICANT: TURLEY, EVA A.
; APPLICANT: SHUWEN, ZHANG
; APPLICANT: ENTWISTLE, JOYCELYN
; TITLE OF INVENTION: HYALURONAN RECEPTOR PROTEIN

```


Db 1355 YVRLKKNVRLABETEKLAENVFLKCKRS 1388

RESULT 13

US-09-723-219-2
; Sequence 2, Application US/09723219
; Patent No. 6391613
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6391613el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/723,219
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/572,191
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-723-219-2

Query Match 17.5%; Score 209; DB 4; Length 1388;

Best Local Similarity 24.8%; Pred. No. 3.6e-08;
Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;

QY 1 QEKYDSMVQSLDY-----TAQESYKA-----LTASEIEDLKLENSLOEKA 43
Db 1118 KNEVFMQRQLHVDNAEDPQPPHPHOTHAKLLETQEQEIEDGRASKTSLEHLV 1177
QY 44 AKAGNAEDVQHQILATESSNQYVRMLDLQ---TKSALKETRIKITVSLQKITDLQ 100
Db 1178 TKLNDREVKNAGELIRMKQREMLRLESQILKXNLWLGQGLDDIK---RQKENSQ 1234
QY 101 N-----OLKQOEEDFRKQ-----LEDEGRKAEKENTTAULTTEINKW 138
Db 1235 NHPDNOQLKQEQESIERLAKSKIVBEMLKWKADLEEVQSALYNKMECLRMMDVEYRT 1294
QY 139 RLVEELYNKTKPFQIOLDAPEVEKQALLNHGAAQOLNIRDSYAKLGHQNLKOKIK 198
Db 1295 QTLSEKAFQEKQLRSLKLEMYERERTSQEMEMLRKQVECLARENGKLVGHNHOKIQ 1354
QY 199 HVKLKDNSOLKSEVSKLRQ---LAKKQOSET 229
Db 1355 YVRLKKNVRLABETEKLAENVFLKCKRS 1388

RESULT 14

US-09-104-324B-4
; Sequence 4, Application US/09104324B
; Patent No. 6232460
; GENERAL INFORMATION:
; APPLICANT: Treci, Ozlem; Sahin, Ugur; Pfreundschuh, Michael
; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,
; TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers In A Sample Of
; TITLE OF INVENTION: No. 6232460mal Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/104,324B
; FILING DATE: 25-June-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/892,702
; FILING DATE: 15-July-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6232460man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5491
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-104-324B-4

Query Match 16.0%; Score 191; DB 3; Length 976;
Best Local Similarity 23.3%; Pred. No. 5.9e-07;
Matches 79; Conservative 53; Mismatches 103; Indels 104; Gaps 12;

QY 1 QEKYDSMVQSLDYVTAQESYKALTAS---EIEDLK-----LENSLOEKAKAG 47
Db 402 EDQKILTMELQKSSLEBEMTKLTNNKEVELBELKVLGKETILLYENKQF-EKIAEEL 460
QY 48 KNAB-----KEITVFLQKITDLQNLQOEEDFRKQLEDEGRKAEKENTTAULTTEINKW 138
Db 461 KGTQEOLIGLQAREKVDHLEIQLTAITTSQYVSKVDLKTLENEKLNKTELTSHC 520
QY 86 -----KEITVFLQKITDLQNLQOEEDFRKQLEDEGRKAEKENTTAULTTEINKW 138
Db 521 NKLSLENKELT---QETSDMTLELNQOEDINNKKQBERMLKQIENLOETETOLRNEI 576
QY 139 RLVEELYNKTKPFQIOLDAFE-----VEKQALLNEHGAQOBL 177
Db 577 EYVRELKQKRDVEKCKLDKSENCNNLRKQVENKQYIELOQENKALKKKGTAEKQL 636
QY 178 N-----KIRDSYA-----KLIGHQNLKQIKHVVVKLKDENSOLK 211
Db 637 NVVEIKVKNLELESQKQKFGITDITYQKEIEDKISENLLLEVEKAKVIADAEVKIQ 696
QY 212 SEVSKLRQ-----LAKKQOSETKLOEELNKVLGI 241
Db 697 KEIDK-RCQHKIABMVALMEKHKHQYDKITEERDSELGL 734

RESULT 15

US-09-310-187A-1
; Sequence 1, Application US/09310187A
; Patent No. 6358751
; GENERAL INFORMATION:
; APPLICANT: Benichou, Gilles
; APPLICANT: Fedoseyeva, Eugenia
; TITLE OF INVENTION: Involvement of Autoantigens in Cardiac
; TITLE OF INVENTION: Graft Rejection
; FILE REFERENCE: UCSF-090
; CURRENT APPLICATION NUMBER: US/09/310,187A
; CURRENT FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-310-187A-1

Query Match 13.6%; Score 162.5; DB 4; Length 1939;
Best Local Similarity 21.1%; Pred. No. 0.00021;
Matches 76; Conservative 50; Mismatches 99; Indels 135; Gaps 12;

QY 8 VOSLEDVTAQFESYKALTASIEDLKLENSLSQEKAAKAGNAEDVQHQILATESNOEY 67
 Db 1209 IDNLRVQKLEKSEFKLEDDV-----TSNMQIIKAKANLEKVSRTL---EQANEY 1261
 QY 68 -----VRMLDLQTSKALKETE-----IKEIT-----VSFLOKITDLOQL 103
 Db 1262 RVKLEEAQSLNDFTTQAKLOTENGELARQLEKEALISQITRGKLSYTTQOMEDLKQOL 1321
 QY 104 KOQBED-----FRKQLEDEBGRKAENTTAELTEBINKWRLLY----- 142
 Db 1322 BEEGKAKNALAHALQSAHRDCDLLREQVEEEMAKELQRLVSKANSEVAQWRTKYETDA 1381
 QY 143 -----EELYNKTKPFOIOL-----DAFEVEKQALLNE-----HGAQ 174
 Db 1382 IORTTELEEAQKLAQRLQDAEBAEAVNAKCSSLEKTKHRLQNEIEDLMVDVRSNAAA 1441
 QY 175 EQLNKIRDSYAKLLGHQNLK-----KLRQLAKKK-----QSETKLQBELNKL 204
 Db 1442 AALDKQKRNFDKILAEWKQKYESOSESSQKEARSLSSTELFKLNAYEESLEHLETFK 1501
 QY 205 DENSQKSEVS-----KLRQLAKKK-----QSETKLQBELNKL 239
 Db 1502 RENKNLQBEISDLTEQLGEGGKNVHELEKVRKQLEVEKLEQSALEEAASLEHEEGKIL 1561

RESULT 16

US-08-938-105-3
 ; Sequence 3, Application US/08938105
 ; Patent No. 6353151
 ; GENERAL INFORMATION:
 ; APPLICANT: Leinwand, Leslie A.
 ; TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sheridan Ross P.C.
 ; STREET: 1700 Lincoln St., Suite 3500
 ; CITY: Denver
 ; STATE: CO
 ; COUNTRY: U.S.A.
 ; ZIP: 80203
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/938,105
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crook, Wainell M.
 ; REGISTRATION NUMBER: 31,071
 ; REFERENCE/DOCKET NUMBER: 3595-4
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (303) 863-9700
 ; TELEFAX: (303) 863-0223
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1886 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

US-08-938-105-3

Query Match 13.5%; Score 161.5; DB 4; Length 1886;
 Best Local Similarity 20.8%; Pred. No. 0.00025;
 Matches 75; Conservative 49; Mismatches 101; Indels 135; Gaps 11;
 QY 8 VOSLEDVTAQFESYKALTASIEDLKLENSLSQEKAAKAGNAEDVQHQILATESNOEY 67
 Db 1156 IDNLRVQKLEKSEFKLEDDV-----TSHMQIIKAKANLEKVSRTL---EQANEY 1208

QY 68 -----VRMLDLQTSKALKETEIKET-----VSFLOKITDLOQL 103
 Db 1209 RVKLEEAQSLNDFTTQAKLOTENGELARQLEKEALINQITRGKLSYTTQOMEDLKQOL 1268
 QY 104 KOQBED-----FRKQLEDEBGRKAENTTAELTEBINKWRLLY----- 142
 Db 1269 BEEGKAKNALAHALQSAHRDCDLLREQVEEEMAKELQRLVSKANSEVAQWRTKYETDA 1328
 QY 143 -----EELYNKTKPFOIOL-----DAFEVEKQALLNE-----HGAQ 174
 Db 1329 IORTTELEEAQKLAQRLQDAEBAEAVNAKCSSLEKTKHRLQNEIEDLMVDVRSNAAA 1388
 QY 175 EQLNKIRDSYAKLLGHQNLK-----KLRQLAKKK-----QSETKLQBELNKL 204
 Db 1389 AALDKQKRNFDKILAEWKQKYESOSESSQKEARSLSSTELFKLNAYEESLEHLETFK 1448
 QY 205 DENSQKSEVS-----KLRQLAKKK-----QSETKLQBELNKL 239
 Db 1449 RENKNLQBEISDLTEQLGEGGKNVHELEKIRKQLEVEKLEQSALEEAASLEHEEGKIL 1508

RESULT 17

US-08-592-126-148
 ; Sequence 148, Application US/08592126
 ; Patent No. 5821091
 ; GENERAL INFORMATION:
 ; APPLICANT: Gregory Dolganov
 ; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
 ; TITLE OF INVENTION: Polypeptides
 ; NUMBER OF SEQUENCES: 151
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Denlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/592,126
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sholtz, Charles K.
 ; REGISTRATION NUMBER: 38,615
 ; REFERENCE/DOCKET NUMBER: 4600-0111
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 324-0880
 ; TELEFAX: (415) 324-0960
 ; INFORMATION FOR SEQ ID NO: 148:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1312 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
 ; US-08-592-126-148

Query Match 13.3%; Score 159; DB 2; Length 1312;
 Best Local Similarity 22.3%; Pred. No. 0.00025;
 Matches 54; Conservative 56; Mismatches 88; Indels 44; Gaps 9;
 QY 2 EKYSMVOSLEDVTAQFESYKALTASIEDLKLENSLSQEKAAKAGNAEDVQHQILATE 61
 Db 796 ERFQMKELKDVVERKIAQ-----QAAKLGIDLDRTVQVQVQKQEKQHK-LDTV 842
 QY 62 SSNOEYVRMLDLQTSKALKETEIKETVSFLQKITDLOQLKOQBEDFRKQLEDEBGRK 121

Db 843 SSKIENRKLIOQOEQ-----IQHLKSTTNELKSEKLOISTNLQ-----RR 884
 QY 122 AEKENTTAELTEINKRWRLLYEELYN---KTRPFQIQLDFAFEVEKQALLNEHGA-----AQ 174
 Db 885 QOLEEQTVLSTEVQS---LYREIKDAKEQVSPLETTLEKFFQOEKEELINKKNTSNKIAQ 941
 QY 175 EOLNKIRDSYAKLLGHONLKQKIKHVVKLKDE-NSOLKSEVSKLRQOLAKKKQSETKLOE 233
 Db 942 DLKINDIKERVKVNIHG- -MKDIENVIQDGKDDYKQKQETELNKVIAQLSECEKHEKINE 999
 QY 234 EL 235
 Db 1000 DM 1001

RESULT 18
 US-08-687-080-51
 ; Sequence 51, Application US/08687080
 ; Patent No. 5965427
 ; GENERAL INFORMATION:
 ; APPLICANT: Gregory Dolganov
 ; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
 ; NUMBER OF SEQUENCES: 175
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/687,080
 FILING DATE: 17-JUL-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/592,126
 FILING DATE: 26-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Sholtz, Charles K.
 REGISTRATION NUMBER: 38,615
 REFERENCE/DOCKET NUMBER: 4600-0111.30
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 51:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1312 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: TRANS. OF RAD50 cDNA (SEQ. 54), NT.
 INDIVIDUAL ISOLATE: 389 TO 4324

US-08-687-080-51
 Query Match 13.3%; Score 159; DB 2; Length 1312;
 Best Local Similarity 22.3%; Pred. No. 0.00025;
 Matches 54; Conservative 56; Mismatches 88; Indels 44; Gaps 9;
 QY 2 EKXDSMVQSLSDVTAQFESYKALTASIEDLKLNSLSQEAAGKAGKNAEDVQHOILATE 61
 Db 796 ERFQMLKDVVERKIAQ-----QAQKLGIDLDRTVQQVQNEKQKHK-LDTV 842
 QY 62 SSNOEYVRMLDLQTKSALKETEITVSLQKITDLQNLKQOEDEFKQLEDEEGRK 121

Db 843 SSKIENRKLIOQOEQ-----IQHLKSTTNELKSEKLOISTNLQ-----RR 884
 QY 122 AEKENTTAELTEINKRWRLLYEELYN---KTRPFQIQLDFAFEVEKQALLNEHGA-----AQ 174
 Db 885 QOLEEQTVLSTEVQS---LYREIKDAKEQVSPLETTLEKFFQOEKEELINKKNTSNKIAQ 941
 QY 175 EOLNKIRDSYAKLLGHONLKQKIKHVVKLKDE-NSOLKSEVSKLRQOLAKKKQSETKLOE 233
 Db 942 DLKINDIKERVKVNIHG- -MKDIENVIQDGKDDYKQKQETELNKVIAQLSECEKHEKINE 999
 QY 234 EL 235
 Db 1000 DM 1001

RESULT 19
 US-09-168-595-148
 ; Sequence 148, Application US/09168595
 ; Patent No. 6555666
 ; GENERAL INFORMATION:
 ; APPLICANT: Gregory Dolganov
 ; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
 ; NUMBER OF SEQUENCES: 151
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Avenue, Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/168,595
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/592,126
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Sholtz, Charles K.
 REGISTRATION NUMBER: 38,615
 REFERENCE/DOCKET NUMBER: 4600-0111
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 148:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1312 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: rad50.pro-translation of SEQ ID NO:54
 INDIVIDUAL ISOLATE: 13.3%; Score 159; DB 4; Length 1312;
 Best Local Similarity 22.3%; Pred. No. 0.00025;
 Matches 54; Conservative 56; Mismatches 88; Indels 44; Gaps 9;

US-09-168-595-148
 Query Match 13.3%; Score 159; DB 4; Length 1312;
 Best Local Similarity 22.3%; Pred. No. 0.00025;
 Matches 54; Conservative 56; Mismatches 88; Indels 44; Gaps 9;
 QY 2 EKXDSMVQSLSDVTAQFESYKALTASIEDLKLNSLSQEAAGKAGKNAEDVQHOILATE 61
 Db 796 ERFQMLKDVVERKIAQ-----QAQKLGIDLDRTVQQVQNEKQKHK-LDTV 842
 QY 62 SSNOEYVRMLDLQTKSALKETEITVSLQKITDLQNLKQOEDEFKQLEDEEGRK 121
 Db 843 SSKIENRKLIOQOEQ-----IQHLKSTTNELKSEKLOISTNLQ-----RR 884

RESULT 23

```

; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2482 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-328-254-6

Query Match      12.7%; Score 152; DB 1; Length 2482;
Best Local Similarity 23.4%; Pred. No. 0.0019;
Matches 67; Conservative 56; Mismatches 89; Indels 74; Gaps 11;

Qy 8 VQSLDVTQAQESYKALTASIEDLKLENSLOEKAAKAGNAEDVQHILATSSNOEY 67
Db 253 ISKLQEDTSAHQNVVAETLSALENKEKELQLNDK-----VETEQAETELKKSNNLL 305
Qy 68 VRMLLDLQTKS-----ALKETEIKETV--SFLQKITDLQNLK---QOE 107
Db 306 EDSLKELQLLSETLSLEKKESSIIISLNKREIBELTQENGTLKEINASLNQEKNNLIQKS 365
Qy 108 EDFRKQLEDEGRKAEXENTTAELTEINKWRL-----YEELYNKTKPFQIQL 156
Db 366 ESFANYID-----EREKSISELSQYKQEKILLQRCETGNAYEDLSQKYAAQEK 418
Qy 157 DAFVEKQALLNEHGAQOE---QLNKIRDSYAKLLGH-----QNLKQKIK 198
Db 419 SKLE-----CLLNECTSLCENKNELEQLKEAFK--EHQFELTKLAFABERNQNLMLELE 472
Qy 199 HV-----VKLDENSQKSEVSKRLCOLAKKQKQSETKLQELNKVL 239
Db 473 TVQQALRSEMTDQNNKSEAGGLKQEIWLKEBQNKQKQEVNDLL 518
```

```

RESULT 25
US-08-353-700-1
; Sequence 1, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,700
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, JANET E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
; US-08-353-700-1

Query Match      12.7%; Score 152; DB 1; Length 3248;
Best Local Similarity 23.4%; Pred. No. 0.0026;
Matches 67; Conservative 56; Mismatches 89; Indels 74; Gaps 11;

Qy 8 VQSLDVTQAQESYKALTASIEDLKLENSLOEKAAKAGNAEDVQHILATSSNOEY 67
Db 885 ISKLQEDTSAHQNVVAETLSALENKEKELQLNDK-----VETEQAETELKKSNNLL 937
Qy 68 VRMLLDLQTKS-----ALKETEIKETV--SFLQKITDLQNLK---QOE 107
Db 938 EDSLKELQLLSETLSLEKKESSIIISLNKREIBELTQENGTLKEINASLNQEKNNLIQKS 997
Qy 108 EDFRKQLEDEGRKAEXENTTAELTEINKWRL-----YEELYNKTKPFQIQL 156
Db 998 ESFANYID-----EREKSISELSQYKQEKILLQRCETGNAYEDLSQKYAAQEK 1050
Qy 157 DAFVEKQALLNEHGAQOE---QLNKIRDSYAKLLGH-----QNLKQKIK 198
Db 1051 SKLE-----CLLNECTSLCENKNELEQLKEAFK--EHQFELTKLAFABERNQNLMLELE 1104
Qy 199 HV-----VKLDENSQKSEVSKRLCOLAKKQKQSETKLQELNKVL 239
Db 1105 TVQQALRSEMTDQNNKSEAGGLKQEIWLKEBQNKQKQEVNDLL 1150

RESULT 26
PCT-US95-16216-1
; Sequence 1, Application PC/TUS9516216
; GENERAL INFORMATION:
; APPLICANT: Yen, Timothy J.
; APPLICANT: Rattner, Jerome B.
; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently
; TITLE OF INVENTION: Expressed Kinetochore Protein, and Methods of Use
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16216
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,700
; FILING DATE: 09-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3248 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
```



```

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-16216-1

Query Match      12.7%; Score 152; DB 5; Length 3248;
Best Local Similarity 23.4%; Pred. No. 0.0026;
Matches 67; Conservative 56; Mismatches 89; Indels 74; Gaps 11;

QY 8 VOSLEDVTAQESYKALTASIEDLKLENSLOEAKAGKNAEDVQHIIATSSNOEY 67
Db 885 ISKLOEDTSRQNVVAETLSALENKEKELQINDK-----VETEQAIEIOLKXSNHLL 937
QY 68 VRMLLDLQTS-----ALKETIKEITV--SFLOKITDLOQLK-----QOE 107
Db 938 EDSLKELQLSETLSLEKKNSSIIINKRIEIBELTQENGTLKEINSLNOKKNLQKS 997
QY 108 EDFKQLEDEGRKAERKENTTAETEEINKWRLL-----YEELYNKTRPFQIOL 156
Db 998 ESFANYID-----EREKSISELSDQYKQEKLIILQRCETGNAYEDLSQYKAAQEK 1050
QY 157 DAFEVEKQALNEHGAQOE-----QINKIRDSYAKLLGH-----QMLKQKIK 198
Db 1051 SKLE-----CLNCTSLCENKNELEQKKEAFK--EHQEFILTKLAFABERNQMLMLE 1104
QY 199 HV-----VKLDENSQKSEVSKRLCOLAKKQSETKLOEBLNKVL 239
Db 1105 TVQALRSEMTDNQNSKSEAGGLKQEIIMTLKEFQNKQKQVNDLL 1150

RESULT 27
US-08-685-871-2
; Sequence 2, Application US/08685871
; Patent No. 6013499
; GENERAL INFORMATION:
; APPLICANT: NARUMIYA, Shuh
; APPLICANT: IWAMATSU, Akihiro
; TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,871
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184102
; FILING DATE: 25-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-262553
; FILING DATE: 14-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18887/845
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 1354 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-685-871-2

Query Match      12.6%; Score 151; DB 3; Length 1354;
Best Local Similarity 24.3%; Pred. No. 0.0011;
Matches 66; Conservative 58; Mismatches 94; Indels 54; Gaps 11;

QY 5 DSMVQSLDVTAQESYKALTASIEDLKLENSLOEAKAGKNAEDVQH-----DL 99
Db 836 EGQMRLEQDQLEAQYFSTLYKTQVKELKE-----IEKNRENKKIOELQNEKETLATOL 892
QY 57 -ILATESNOEYVRMLLDLO-----TKSAL-----KETEIKEITVVSFLOKIT----- 99
Db 893 DLAEKASQLAGLEBQVFELOESKKAASNRROEITDKOHTVSRLEEANSMLTKDI 952
QY 100 QNOLKQOEEDFRKQLEDEGRKAEKE-----NTTAELTEEINKWRLLLYEELYNK-----T 149
Db 953 EILRRENEBELTKMKAAEEYKLEKEBEISNLKAAFEKNINTERTLTATQAVNKLAEINMR 1012
QY 150 KPFOIQLDAFEVKQALL-----NEHGAQOEOLKIRDSYAKLLGHONLQKIKHVVVKLDE 206
Db 1013 KDPKI--DRKKANTQDLRKKENRKLQELNOREKFNQMV-----VKHOKELNDM 1062
QY 207 NSOLKSEV---SKRLCOLAKKQSETKLOEBL 235
Db 1063 QAQLVECAHRNELQMLASKESDIEQIRAKL 1094

RESULT 28
US-09-595-684B-31
; Sequence 31, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; FILE REFERENCE: cyto036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 2662
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-31

Query Match      12.5%; Score 149.5; DB 4; Length 2662;
Best Local Similarity 25.5%; Pred. No. 0.0032;
Matches 73; Conservative 48; Mismatches 90; Indels 75; Gaps 14;

QY 1 QEYDSMVQSLDVTAQESYKA-----ITASEIEDLKLENSL-----QEKAAK 45
Db 1049 QEK-NELQOMLESVIAEKEQLKTDLKNIEWTIENQELALLGDELKQOEIVAEQKNHA 1107
QY 46 AGQAE-----DVQHQILATESNQYVRMLLDLOPKSALKETEIKEITVVSFLOKI 96
Db 1108 IKKEGELSRCTDRLAEEVEKLEKSKQQLQKQQLNVOEEMSEMOKKINE-----I 1159
QY 97 TDLQNLKQOEEDFRKQLEDEGRKAEKENTTAELTEEINKWRLLLYEELYNKTRPFQIOL 156
Db 1160 ENLKNELKNKELTY-EHMTERLELELAQKUNENVEEYKSIKTKRKVLKEL---QKSPETER 1215
QY 157 DAF-----EVEKQAL-----LNEHGAQOEOLKIRDSY-----AKLLGHQNLKQ 195
Db 1216 DHLRGVIREIEATGLQTKELKTAHILKEH---QETIDELRRSVSEKTAQIINTQDLEK 1272
```

```
Qy 196 ---KXHVVKLDNSQKSEVSKLRCLAKKQSETKLOELNKV 238
Db 1273 SHTKLOEIPVHBEQELPNV-----KQVSET--QETMNEL 1307

RESULT 29
US-09-103-664A-2
; Sequence 2, Application US/09103664A
; Patent No. 6458358
; GENERAL INFORMATION:
; APPLICANT: University of Kentucky Research Foundation
; APPLICANT: Timoney, John
; APPLICANT: Ariushin, Sergey
; TITLE OF INVENTION: Compounds Encoding the Protective M-Like Protein of Streptococcus
; FILE REFERENCE: 50229-212
; CURRENT APPLICATION NUMBER: US/09/103,664A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Streptococcus equi
US-09-103-664A-2

Query Match 12.2%; Score 146; DB 4; Length 534;
Best Local Similarity 25.4%; Pred. No. 0.00088;
Matches 73; Conservative 50; Mismatches 102; Indels 62; Gaps 13;

Qy 1 QKYSVMQSLSDVTAQPSY-KALTAS--ETEDIKLENSSLQEKAAKAGK-----NA 50
Db 174 KEKEQELTKAKDERQALTESFNKTLRSSTKEYNKLKTELAKKEKAQKTKELADKLSNA 233
Qy 51 E-----DVQHOILATESSSNOEYVRMLDLQTKSALKETBIKITVSFLOKITDL 99
Db 234 EASRDKAFAVSKDLADKLSSAASRDKAPAVSKDLADKLAAKTAEEKL-----MENVGSL 289
Qy 100 QNQLKQBEDFRKQLEDEBGRKAEKENTTAETLTEINKWRLLYBELYNKTKPFQIOLDAP 159
Db 290 DRLVESAKREMAQKLAIEIDQLTADKADAKADAELAAANDTASLOTBL-EKAKT----- 340
Qy 160 EVEKQALLNEHGAQ-EQANKIRDSYAKLL--GHQNL-----KQI---KHVVKLDENSQ 209
Db 341 ELAVSERLIESGRHIAELQKQDASDKALVESQANVAELEKOKAKASDAKVAELEKEVEA 400
Qy 210 LKSEVSKLRCLAKKQKQ-----ETKLQ-----BELNKV 238
Db 401 AKAEVADLKAQLAKKEBELEAVKKEALEAKTEELKKAHAEELSGL 447

RESULT 30
US-09-914-259-11
; Sequence 11, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 3878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-259-11

Query Match 12.2%; Score 146; DB 4; Length 3878;
```

```
Best Local Similarity 26.8%; Pred. No. 0.0092;
Matches 79; Conservative 47; Mismatches 75; Indels 94; Gaps 19;

Qy 2 EKYSVMQSL-----EDVTAQ--FESYKALTASEIEDLKLENS----- 38
Db 290 EQDHLLEDYQKKKEDFTMOISFLQEKIKVYEMEQDKKVENSKKEIQEKETIIEELNTK 349
Qy 39 -----LQEKAAKAGNAEDVOHOILATESSSNOEYVRMLDLQTKSALKKE-----TE 84
Db 350 IIEEKKTLLEKDKLTADKLLGLQEQIV--QXNOEIKNMKLEL-TNSKQKERSSEE 405
Qy 85 IKETI--TVSFLQKITDILQNLKQOEBDFRKQLEDEBGRKAEKENTTAETLTEEINKWRLLY 142
Db 406 IKQLMGIVBELOKENHKDSQF---ETDIVQMEQETQKLEQ--LRAEL----- 449
Qy 143 BELYNKTKPFQIOLDAPFEVEKQALLNEHGAQEQOL-----NKIRDSYAKLLGHQNL 193
Db 450 DEMYGOQ---IVQM-----KQELIRQHMAQMEEMKTRHKGEMENALR-SYSNITVNE- 498
Qy 194 KQIKHV-----VKLDENSQ---LKSEVSKL---RCOLAKKQSETKLOEELN 236
Db 499 QIKLMNVAINELNFKLODTNSQEKLEELGLILEEKCALQORQED---LVEELS 550

Search completed: December 16, 2003, 06:09:56
Job time : 23 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:07:28 ; Search time 21 Seconds
(without alignments)
1108.230 Million cell updates/sec

Title: US-09-978-309A-74

Perfect score: 1195

Sequence: 1 QEKYDSMVQSLDVTAFES.....KKQSETKLOELNKVLGIK 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database :

PIR_76:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1193	99.8	725	1 JC5016	hyaluronan recepto
2	914	76.5	631	2 JC4298	hyaluronan recepto
3	193.5	16.2	955	2 S24348	myosin heavy chain
4	193	16.2	1388	2 T30335	KLP2 protein - Afr
5	191	16.0	1410	1 A57013	early endosome ant
6	190	15.9	1690	2 T13030	microtubule bindin
7	187.5	15.7	741	2 S39082	myosin heavy chain
8	185.5	15.5	1790	2 S67593	transport protein
9	182.5	15.3	1938	2 A52993	skeletal myosin he
10	181.5	15.2	1940	1 A24922	myosin heavy chain
11	179.5	15.0	936	2 S39083	myosin heavy chain
12	179	15.0	697	2 T07111	MAR binding filame
13	178.5	14.9	1940	1 S04090	myosin heavy chain
14	177.5	14.8	845	1 J00178	myosin heavy chain
15	176.5	14.8	1937	2 A48176	synaptonemal compl
16	176.5	14.8	1937	2 A38055	myosin heavy chain
17	176.5	14.8	1940	2 A23320	myosin heavy chain
18	175.5	14.7	876	2 A23767	myosin heavy chain
19	175	14.6	993	2 S49461	synaptonemal compl
20	172.5	14.4	439	2 A06068	paramyosin - fluke
21	172.5	14.4	864	2 B90395	purine NTPase [imp
22	170	14.2	1388	2 S74245	serine/threonine-s
23	169	14.1	764	2 I51302	myosin heavy chain
24	169	14.1	1931	2 A59234	myosin heavy chain
25	167	14.0	2245	2 T18278	slow myosin heavy
26	166.5	13.9	1039	2 S18199	myosin heavy chain
27	166	13.9	1300	2 I53799	myosin heavy chain
28	166	13.9	1356	2 S32763	kinectin 1 - huma
29	165	13.8	3187	2 JC5837	364K Golgi complex

30	162.5	13.6	1164	2	T24806	hypothetical prote
31	162.5	13.6	1938	2	I49464	alpha cardiac myos
32	162	13.6	1935	2	A59286	myosin heavy chain
33	161.5	13.5	911	2	S51441	hypothetical prote
34	161.5	13.5	1938	1	S06005	myosin alpha heavy
35	161.5	13.5	1939	2	I48175	myosin heavy chain
36	161	13.5	746	2	T47237	myosin II heavy ch
37	161	13.5	1935	1	A37102	myosin beta heavy
38	160.5	13.4	4574	2	G02520	plectin - human
39	160	13.4	1679	2	S48385	hypothetical prote
40	160	13.4	1738	2	T14867	interaptin - slime
41	160	13.4	1935	1	S06006	myosin beta heavy
42	160	13.4	1979	2	S03186	myosin heavy chain
43	159.5	13.3	1597	2	S68420	ctron - mouse
44	158	13.2	978	2	A70387	conserved hypothet
45	158	13.2	1156	2	B70356	chromosome assembl
46	158	13.2	1203	2	B55094	chromosomal protei
47	158	13.2	1974	2	T16703	hypothetical prote
48	158	13.2	2139	2	T18296	myosin heavy chain
49	157.5	13.2	946	2	S28061	SCP1 protein - rat
50	156.5	13.1	886	2	H69378	conserved hypothet

ALIGNMENTS

RESULT 1

JC5016

hyaluronan receptor - human

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999

C:Accession: JC5016

R:Wang, C.; Entwistle, J.; Hou, G.; Li, Q.; Turley, E.A.

Gene 174, 299-306, 1996

A:Title: The characterization of a human RHAMM cDNA: Conservation of the hyaluronan-bi

A:Reference number: JC5016; MUID:97045829; PMID:8890751

A:Contents: breast

A:Accession: JC5016

A:Molecule type: mRNA

A:Residues: 1-725 <WAN>

A:Cross-references: GB:U29343

A:Note: it is uncertain whether Met-1 or Met-196 is the initiator

C:Comment: This receptor regulates focal adhesion turnover, and regulates human breast

C:Genetics:

A:Gene: GDB:HMWR; RHAMM

A:Cross-references: GDB:683209; OMIM:600936

A:Map position: 5q33.2-5qter

C:Superfamily: hyaluronan receptor

Query Match 99.8%; Score 1193; DB 1; Length 725;

Best Local Similarity 99.6%; Pred. No. 1.9e-53;

Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QEKYDSMVQSLDVTAFESYKALTASEIEDLKLENSLOEKAAGKAGNAEDVQHILAT 60

Db 443 QEKYDSMVQSLDVTAFESYKALTASEIEDLKLENSLOEKAAGKAGNAEDVQHILAT 502

Qy 61 ESSNQEVYRMLLDIQTSALKETIKETIVSFLOKITDQLNQLKQBEDFRKQLEDEGR 120

Db 503 ESSNQEVYRMLLDIQTSALKETIKETIVSFLOKITDQLNQLKQBEDFRKQLEDEGR 562

Qy 121 KAEKENTABETINKRLLYELNKKTPFQQLDAFEVEKQALLNEHGAAGQQLNKI 180

Db 563 KAEKENTABETINKRLLYELNKKTPFQQLDAFEVEKQALLNEHGAAGQQLNKI 622

Qy 181 RDSYAKLLGHONLKQKTKHVVKLDENSOLKSEVSKLRCOLAKKQSETKLOELNKVLG 240

Db 623 RDSYAKLLGHONLKQKTKHVVKLDENSOLKSEVSKLRCOLAKKQSETKLOELNKVLG 682

Qy 241 IK 242

Db 683 IK 684

```

RESULT 2
Jc4298
A:Species: Mus musculus (house mouse)
C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 01-Dec-2000
C:Accession: Jc4298; A42925; A41923; S21586
R:Ensemble: J.; Zhang, S.; Yang, B.; Wong, C.; Li, Q.; Hall, C.L.; A, J.; Mowat, M.; G.
Gene 163, 233-238, 1995
A:Title: Characterization of the murine gene encoding the hyaluronan receptor RHAMM.
A:Reference number: Jc4298; MUID:96011639; PMID:7590272
A:Accession: Jc4298
A:Molecule type: mRNA
A:Residues: 1-631 <ENT>
A:Cross-references: EMBL:X64550
A:Experimental source: 3T3 fibroblast
R:Hardwick, C.
J. Cell Biol. 118, 753, 1992
A:Reference number: A42925; MUID:92348516; PMID:1639856
A:Contents: erratum
A:Accession: A42925
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 155-376, 'S', 378-504, 'E', 506-631 <HAR>
A:Cross-references: GB:X64550
A:Note: authors translated the codon AGT for residue 377 as Thr and CTA for residue 507
R:Hardwick, C.; Hoare, K.; Owens, R.; Hohn, H.P.; Hook, M.; Moore, D.; Cripps, V.; Auste
J. Cell Biol. 117, 1343-1350, 1992
A:Title: Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motil
A:Reference number: A41923; MUID:92299690; PMID:1376732
A:Accession: A41923
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 155-376, 'S', 378-504, 'E', 506-507, 'I', 508-630 <HA2>
A:Cross-references: GB:X64550
A:Note: This sequence has been corrected in reference A42925
C:Comment: This protein regulates cell motility and transformation, and focal adhesion o
C:Genetics:
A:Gene: Rhamm
A:Introns: 20/3; 54/2; 79/2; 137/3; 187/3; 259/2; 382/2; 431/2; 482/2; 515/3; 574/3; 625
C:Superfamily: hyaluronan receptor
C:Keywords: glycoprotein; receptor
F:260-382/Region: 21 residue repeats
F:516-574/Region: hyaluronan binding #status predicted
F:575-625/Region: hyaluronan binding #status predicted
F:116,283,304,325,346,367,398,438,619/Binding site: carbohydrate (Asn) (covalent) #statu
Query Match 76.5%; Score 914; DB 2; Length 631;
Best Local Similarity 76.0%; Pred. No. 2.1e-39;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;
QY 1 QKYSVMQSLDVTQAFESYKALTASIEDLKLENSLOEKAAGKNAEDVQHOILATESNOEY 60
DB 363 QKYNDAQSLRDVTAQLESYKSLTKETEDLKLENTLOEKVMAEKSVEDVQOILTA 422
QY 61 ESSNOEYVRMLDLDTQKALKETEIKETVSLQKITDLOKQOEDDFRQLEDEGR 120
DB 423 ESTNOEYARWVDLQNRSTLKEEIEKITSSPLEKITDLOKQOEDDFRQLEDEGR 482
QY 121 KAEKNTTAEITPEINKRLLYELLYNKKPQIQLDAFEVEKQALLNEHGAQOQLNKI 180
DB 483 TAEKENVMTLWEINKRLLYDELYEYKTPQQQLDAFEAKQALLNEHGTQOQLNKI 542
QY 181 RDSYAKLGHQNLKOKIKHVKVLDKENSOLKSEVSKLRQALAKKQSETKLQEEINKVLG 240
DB 543 RDSYAKLGHQNLKOKIKHVKVLDKENSOLKSEVSKLRQALAKKQSETKLQEEINKVLG 240
QY 241 IK 242
DB 603 IR 604
RESULT 3

```

```

S24348
myosin heavy chain, embryonic and adult skeletal muscle (clone Cemb2) - chicken (frag
C:Species: Gallus gallus (chicken)
C:Date: 03-Feb-1994 #sequence_revision 06-Sep-1996 #text_change 13-Feb-1998
C:Accession: S24348
R:Moore, L.A.; Arrizubieta, M.J.; Tidman, W.E.; Herman, L.A.; Bandman, E.
J. Mol. Biol. 225, 1143-1151, 1992
A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isofo
A:Reference number: S24348; MUID:92309413; PMID:1377278
A:Accession: S24348
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-955 <MOO>
A:Cross-references: EMBL:M74085
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: ATP; coiled coil; muscle contraction; skeletal muscle; thick filament
Query Match 16.2%; Score 193.5; DB 2; Length 955;
Best Local Similarity 21.3%; Pred. No. 0.0086;
Matches 76; Conservative 53; Mismatches 106; Indels 121; Gaps 9;
QY 8 VQSLDVTQAFESYKALTASIEDLKLENSLOEKAAGKNAEDVQHOILATESNOEY 67
DB 226 IDNLRVQKLEKSELSKNEIDDLASNMESVSKAKANLEKMCRTLEQLSKIKSEEH 285
QY 68 VRMLDLQTKSALKETEIKETVSLQKITDLOKQOEDDFRQLEDEGR 106
DB 286 QRMINDLSTQARLQTEGEYSRQVEEKDALISQLSRGQAFQOIEELKRLHEBEIKAK 345
QY 107 -----BED---FRQLEDEGRKAKENTTAELTEINKWRLY-----BEL 145
DB 346 NALAHALQSAKHDCDLREQYEEQKAGELQALSKANSEVAQWRTKYETDAIQRTTEL 405
QY 146 YNKKPFOQLD-----AFVEKQALLNE-----HGAQOQLNKIR 181
DB 406 EAKKLAQRLODAEEHVEAVNAKASLEKTKQLQNEVEDLMDVDEANACARLDKQ 465
QY 182 DSYAKLGHQNLK-----KKO-----SETKLOEINKVLGK 242
DB 466 KNFDKILAEKQKYEETQAELEAQKESRSLSLTELKMKNAYESLDHLETLKRNKQLQ 525
QY 212 SEVSKLRQALAK-----KKO-----SETKLOEINKVLGK 242
DB 526 QEISDLTEQIAEGGKAIHELEKVKKQIEQEKSEIQAALEAEAEASLEHEEGKILRLQ 581
RESULT 4
TJ30335
KLP2 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 02-Jun-2000
C:Accession: TJ30335
R:Boletti, H.; Karsenti, E.; Vernos, I.
Cell 84, 49-59, 1996
A:Title: Xklp2, a new Xenopus centrosomal kinesin-like protein required for centrosome
A:Reference number: Z20827; MUID:96140639; PMID:8548825
A:Accession: TJ30335
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1388 <BOL>
A:Cross-references: EMBL:X94082; MUID:g1129172; PID:e213754; PIDN:CAA63826.1
C:Genetics:
A:Gene: klp2
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
Query Match 16.2%; Score 193; DB 2; Length 1388;
Best Local Similarity 22.9%; Pred. No. 0.013;
Matches 71; Conservative 51; Mismatches 92; Indels 96; Gaps 7;
QY 9 QSLDVTQAFESYKALTASIEDLKLENSLOEKAAGKNAEDVQHOILATESNOEY 67
DB 1079 EELEKLTAEAFNKQALHTHEKELVEKEQISELTNQKLTDLLEISREQEKIRPASSNS 1138

```


RESULT 8
S67593
transport protein US01 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D2552; protein YDL058W
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence.revision 12-Jul-1996 #text_change 21-Jul-2000
C;Accession: S67593; A38455; S30782
R;Blöcker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67587
A;Accession: S67593
A;Molecule type: DNA
A;Residues: 1-1790 <BLO>
A;Cross-references: EMBL:Z74106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058W
A;Experimental source: strain S288C
R;Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
J. Cell Biol. 113, 245-260, 1991
A;Title: A cytoskeleton-related gene, US01, is required for intracellular protein transport
A;Reference number: A38455; MUID:91185402; PMID:2010462
A;Accession: A38455
A;Molecule type: DNA
A;Residues: 1-389,'TA',392-724,'S',726-1790 <NAK>
A;Cross-references: GB:X54378; NID:g4777; PID:CA338253.1; PID:g4778
A;Note: the authors translated the codon ACT for residue 768 as Ile
R;Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
submitted to the EMBL Data Library, February 1993

QY 107 -----BED-----FRKQLEDEBGRKAEKENTTAELTEINKWRLLY----- 142
 Db 1326 IYAKSALAHALOSARHDCDLLEQYEEBGRKAEKENTTAELTEINKWRLLY----- 1385
 QY 143 -BELYNKTKPFOQLD-----AFEVEKQALLNE-----HGAQAEOL 177
 Db 1386 TEELBEAKKLAQRLQDAEHEVAVNAKASLEKTKQRLQNEVEDLMIDVETNAACAAL 1445
 QY 178 NKIRDSYAKLLG-----HONLQK-----KIKHVK-----LKBN 207
 Db 1446 DKKQNFQKILAEVHKHYEETHAELEASQKESRSLSTEVFKVKNAYESLDQLETKREN 1505
 QY 208 SOLKSEVSKRLCOLAK-----KKO-----SETKLOEELNKVIGIK 242
 Db 1506 KNLQOEISDLTEQIAGGKRTHLEKVKQVQEKSELOALEEASLEHBEKILRIQ 1565
 RESULT 10
 A24922
 myosin heavy chain, skeletal muscle, embryonic - rat
 N:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 19-Apr-2002
 C:Accession: A24922; A22538; B24263
 R:Strehler, E.E.; Strehler-Page, M.A.; Perriard, J.C.; Periasamy, M.; Nadal-Ginard, B.
 J. Mol. Biol. 190, 291-317, 1986
 A:Title: Complete nucleotide and encoded amino acid sequence of a mammalian myosin heavy
 A:Reference number: A24922; MUID:870603988; PMID:3783701
 A:Accession: A24922
 A:Molecule type: DNA
 A:Residues: 1-1940 <STR>
 A:Cross-references: GB:X05004; NID:G56658; PIDN:CAA27817.1; PID:G1619328
 R:Strehler, E.E.; Mahdavi, V.; Periasamy, M.; Nadal-Ginard, B.
 J. Biol. Chem. 260, 468-471, 1985
 A:Title: Intron positions are conserved in the 5' end region of myosin heavy-chain genes
 A:Reference number: A22538; MUID:85080119; PMID:2981212
 A:Accession: A22538
 A:Molecule type: DNA
 A:Residues: 1-168 <ST2>
 A:Cross-references: GB:L00370; GB:M10135; NID:G205580; PIDN:AAA41655.1; PID:G554476
 R:Periasamy, M.; Wydro, R.M.; Strehler-Page, M.A.; Strehler, E.E.; Nadal-Ginard, B.
 J. Biol. Chem. 260, 15856-15862, 1985
 A:Title: Characterization of cDNA and genomic sequences corresponding to an embryonic my
 A:Reference number: A24263; MUID:86059474; PMID:2999140
 A:Accession: B24263
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1358-1490, 'G' <PPR>
 A:Cross-references: GB:K03469; NID:G205573; PIDN:AAA41652.1; PID:G205574
 A:Experimental source: clone pMC-72
 C:Genetics:
 A:Introns: 68/3; 116/3; 169/1
 A>Note: the list of intron positions may be incomplete
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; hydroxylase; methylated amino acid; muscle co
 F:89-767/Domain: myosin motor domain homology <MMOT>
 F:179-186/Region: nucleotide-binding motif A (P-loop)
 F:549-586/Region: actin binding #status predicted
 F:656-678/Region: actin binding #status predicted
 F:840-1940/Domain: coiled coil #status predicted
 F:840-1280/Region: S2
 F:1281-1940/Region: light meromyosin
 F:130/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
 F:185/Binding site: ATP (Lys) #status predicted
 F:696,706/Active site: Cys #status predicted
 Query Match 15.2%; Score 181.5; DB 1; Length 1940;
 Best Local Similarity 19.4%; Pred. No. 0.073;
 Matches 69; Conservative 62; Mismatches 104; Indels 121; Gaps 7;
 QY 8 VQSLDVTAFESYKALTASIEDLKLENSLQEKAKAGKNAEDVQHOILATESNOEY 67
 Db 1208 IDNLRVQKLEKSEFKLEIDDLSSSVESVSKANLEKICRTLEDQLEARGKNEET 1267

QY 68 VRMLDLOTKSALEKTEIKRIT-----VSFLQKITDLOLQKOEED- 109
 Db 1268 QRSLSLTQKSRLOTEAGELSRQLEEKESIVSOLSRKQAFQOIEELKRLQLEENKAK 1327
 QY 110 -----FRKQLEDEBGRKAEKENTTAELTEINKWRLLY----- 143
 Db 1328 NALAHALQSSRHDCDLLEQYEEBGRKAEKENTTAELTEINKWRLLY----- 1387
 QY 144 -----ELYN-----KTK----- 150
 Db 1388 BEAKKLAQRLQDSSEQVEAVNAKASLEKTKQRLQGEVEDLMVDVERANSALALDKKQ 1447
 QY 151 -----PFOQLDAAFEVEKQALLNEHGAQAEOLNKTRDSYAKLLG-----HONLK 194
 Db 1448 RNFQKVLAEWTKCEESQAELEAKESRSLSSTELFKLNAYEALDQLETVKRENKLE 1507
 QY 195 QKIKHVVKLDKENSOLKSEVSKRLCOLAKK-----QSETKLOEELNKVIGIK 242
 Db 1508 QEIADLTQIAGGKSTHELEKSKQMELEKADIQMALEEAALEHBEAKILRIQ 1563
 RESULT 11
 S39083
 myosin heavy chain, neonatal [similarity] - chicken (fragment)
 C:Species: Gallus gallus (chicken)
 C:Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text_change 20-Jun-2000
 C:Accession: S39083; S24350; A26821
 R:Moore, L.A.; Arrizubieta, M.J.; Tidymann, W.E.; Herman, L.A.; Bandman, E.
 submitted to the EMBL Data Library, August 1991
 A:Description: Analysis of the chicken fast myosin heavy chain family: Localization of
 A:Reference number: S39081
 A:Accession: S39083
 A:Molecule type: mRNA
 A:Residues: 1-936 <MOO1>
 A:Cross-references: EMBL:M74087
 R:Moore, L.A.; Arrizubieta, M.J.; Tidymann, W.E.; Herman, L.A.; Bandman, E.
 J. Mol. Biol. 225, 1143-1151, 1992
 A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isofo
 A:Reference number: S24348; MUID:92309413; PMID:1377278
 A:Accession: S24350
 A:Molecule type: mRNA
 A:Residues: 3-466, 'Q', 468-641, 'R', 643-936 <MOO2>
 A:Cross-references: EMBL:M74087
 R:Moore, L.A.; Arrizubieta, M.J.; Tidymann, W.E.; Herman, L.A.; Bandman, E.
 J. Mol. Biol. 225, 1143-1151, 1992
 A:Title: Genomic clones encoding chicken myosin heavy-chain genes.
 A:Reference number: A26821; MUID:87217964; PMID:3034534
 A:Accession: A26821
 A:Molecule type: DNA
 A:Residues: 'P', 856-936 <MOR>
 A:Cross-references: GB:M16557; NID:G212371; PIDN:AAA4970.1; PID:G212372
 C:Genetics:
 A:Introns: 886/3
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; muscle; skeletal muscle
 Query Match 15.0%; Score 179.5; DB 2; Length 936;
 Best Local Similarity 21.1%; Pred. No. 0.043;
 Matches 75; Conservative 55; Mismatches 105; Indels 121; Gaps 11;
 QY 8 VQSLDVTAFESYKALTASIEDLKLENSLQEKAKAGKNAEDVQHOILATESNOEY 67
 Db 208 IDNLRVQKLEKSEFKLEIDDLSSSVESVSKANLEKICRTLEDQLEARGKNEET 267
 QY 68 VRMLD-----LOTKSA-----LKETEIKETVS-----FLQKITDLOLQKOEED- 106
 Db 268 QRMINDVNAQARLOTEGYSYRQVEEKDALISQSRGKQAFQOIEELKRLQLEENKAK 327
 QY 107 -----EED-----FRKQLEDEBGRKAEKENTTAELTEINKWRLLY-----EEL 145
 Db 328 NALAHALQSSRHDCDLLEQYEEBGRKAEKENTTAELTEINKWRLLY-----EEL 145

QY 146 YNKTFFQIQD-----AFVEKQALLNE-----HGAQEQNLKIR 181
 Db 388 EAQKWLQRLQDAEHEVAVNSKASLETKORLQNEVEDLMIDVERANSACALDKQ 447
 QY 182 DSAKLLGHQNLK-----QKIKHVVLKLDNSQLK 211
 Db 448 KNFDKILSEWKQKYEQTQAELEASQKESRSLTELFKMKNAVEESLDHLETKRENKILQ 507
 QY 212 SEVSKLRCOLAK-----KKQ-----SETKLOEELNKVLGIX 242
 Db 508 QEISDLTEQIAEGGKAHELEKVKVKQIOEKSQELQSLAEASLSHERGKILRLQ 563
 RESULT 12
 T07111
 MAR binding filament-like protein 1 - tomato
 N:Alternate names: NFPI protein
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
 C:Accession: T07111
 R:Meier, I.; Phelan, T.; Gruissem, W.; Spiker, S.; Schneider, D.
 Plant Cell 8, 2105-2115, 1996
 A:Title: NFPI a novel plant filament-like protein with affinity for matrix attachment re
 A:Reference number: Z15928; MUID:97112038; PMID:8953774
 A:Accession: T07111
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-697 <MEI>
 A:Cross-references: EMBL:X07861; NID:gl771157; PIDN:CAA69181.1; PID:gl771158
 A:Experimental source: strain VFNT Cherry; young fruit
 C:Genetics:
 A:Gene: NFPI
 C:Superfamily: hyaluronan receptor
 C:Keywords: DNA binding, nuclear matrix

Query Match 15.0%; Score 179; DB 2; Length 697;
 Best Local Similarity 25.6%; Pred. No. 0.034;
 Matches 72; Conservative 57; Mismatches 96; Indels 56; Gaps 12;

QY 1 QEKYDSWQSLQEDVIAQFYSYKALTAS---ETEDLKLENSIQEAKAKAGNAEDVOH-- 55
 Db 427 QSELSNRSREVSDDITVQLEQLRSLKREYSKLQME---LEETRASLQRNIDETKSS 483
 QY 56 QILATE-----SSNOEYVRLMDLQTKGALKETEIKELTVSPLOKITDLOLQKQ 106
 Db 484 ELLAELTTKLLKKTNEEMTMSDELVAENRDSLOTET-VNYKKREHTRNELKOE 542
 QY 107 -----EDFRKQLEDEBGRKAKEKNTTAETLFEINKWRL-LYEEL 145
 Db 543 KTIVRTLEELKFLQSITREKELRKSLEDE---LEK---ATESLDELNRNVLAAEEL 595
 QY 146 YNKTFFQIQDFAFEVKQALLNEHGAQEQNLKIRDSYAKLL-----GHQNLKQIKHV 201
 Db 596 ELATSNLSLEDERVHRQSVSEQKQISQAEQNLDEAHSLVNKLKGERSELEKRAK--- 652
 QY 202 KLKDNSQLKSVSKLRCQAKKQKQSETKLQBELNKVLGIX 242
 Db 653 KLEDEMAAAKGEILRLRSQINSVK---APVEDEKVVAGEK 690

RESULT 13
 S04090
 myosin heavy chain 3, skeletal muscle, embryonic - human
 N:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
 C:Accession: S04090; S06146; S05442; S12460; S09333; A35082
 R:Eller, M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Rubinstein, N.A.; Kelly, A.M.
 Nucleic Acids Res. 17, 3591-3592, 1989
 A:Title: Nucleotide sequence of full length human embryonic myosin heavy chain cDNA.
 A:Reference number: S04090; MUID:89263803; PMID:2726495
 A:Accession: S04090
 A:Status: translation not shown

A:Molecule type: mRNA
 A:Residues: 1-1940 <ELL>
 A:Cross-references: EMBL:X13988; NID:G34843; PIDN:CAA32167.1; PID:G34844
 R:Eller, M.; Stedman, H.H.; Sylvester, J.E.; Fertels, S.H.; Wu, Q.L.; Raychowdhury, M.
 FEBS Lett. 256, 21-28, 1989
 A:Title: Human embryonic myosin heavy chain cDNA. Interspecies sequence conservation o
 A:Reference number: S06146; MUID:90033298; PMID:2806546
 A:Accession: S06146
 A:Molecule type: mRNA
 A:Residues: 774-1662, 'QT', 1665-1940 <EL2>
 A:Cross-references: EMBL:X13100; NID:G31143; PIDN:CAA31492.1; PID:G31144
 R:Karsch-Mizrachi, I.; Travis, M.; Blau, H.; Levinwand, L.A.
 Nucleic Acids Res. 17, 6167-6179, 1989
 A:Title: Expression and DNA sequence analysis of a human embryonic skeletal muscle myo
 A:Reference number: S05442; MUID:89366648; PMID:2771643
 A:Accession: S05442
 A:Molecule type: DNA
 A:Residues: 856-1390, 'KK', 1393-1940 <KAR>
 A:Cross-references: EMBL:X15696; NID:G36504; PIDN:CAA33731.1; PID:G1335313
 R:Stedman, H.H.; Eller, M.; Jullian, E.H.; Fertels, S.H.; Sarkar, S.; Sylvester, J.E.;
 J. Biol. Chem. 265, 3568-3576, 1990
 A:Title: The human embryonic myosin heavy chain. Complete primary structure reveals ev
 A:Reference number: A35082; MUID:90154023; PMID:2303463
 A:Contents: annotation; chromosomal assignment
 R:Bober, E.
 Submitted to the EMBL Data Library, January 1989
 A:Reference number: S12458
 A:Accession: S12458
 A:Molecule type: mRNA
 A:Residues: 856-1330, 'G', 1332-1390, 'KK', 1393-1607, 'RA', 1610-1940 <BOB>
 A:Cross-references: EMBL:X51593; NID:G29463; PIDN:CAA35942.1; PID:G29464
 A:Experimental source: Clone gtMHC-E
 R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.
 Eur. J. Biochem. 189, 55-65, 1990
 A:Title: Identification of three developmentally controlled isoforms of human myosin h
 A:Reference number: S09331; MUID:90235862; PMID:1691980
 A:Accession: S09331
 A:Molecule type: mRNA
 A:Residues: 856-901, 'X', 903-971, 'X', 973-1041, 'X', 1043-1111, 'X', 1113-1181, 'X', 1183-1251
 I, 'X', 1673-1741, 'X', 1743-1811, 'X', 1813-1881, 'X', 1883-1940 <BOV>
 A:Cross-references: EMBL:X51593
 C:Genetics:
 A:Gene: GDB:MYH3
 A:Cross-references: GDB:119443; ONIM:160720
 A:Map position: 17p13.1-17p13.1
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle
 F:89-767/Domain: myosin motor domain homology <MMOT>
 F:179-186/Region: nucleotide-binding motif A (P-loop)
 F:549-586/Region: actin binding #status predicted
 F:656-678/Region: actin binding #status predicted
 F:840-1940/Domain: coiled coil #status predicted <COI>
 F:840-1280/Region: S2
 F:1281-1940/Region: light meromyosin
 F:130/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
 F:185/Binding site: ATP (Lys) #status predicted
 F:696,706/Active site: Cys #status predicted

Query Match 14.9%; Score 178.5; DB 1; Length 1940;
 Best Local Similarity 19.1%; Pred. No. 0.1;
 Matches 68; Conservative 62; Mismatches 105; Indels 121; Gaps 7;

QY 8 VQSLQEDVIAQFYSYKALTASIEDLKLENSIQEAKAKAGNAEDVOHILATSSNOEY 67
 Db 1208 IDNLQVRVKQKLEKSEFKLEIDDLSSSMESVSKANLEKICRTLEDQLSEARGKNEEI 1267
 QY 68 VRMLDLQTKGALKETEIKET-----VSFLQKITDLOLQKQESD- 109
 Db 1268 QRSLSLTTQKSLRQLTQAGELSRQLEKEISVLSQSKQAFTQCTBELKQLEENKAK 1327
 QY 110 -----FRKQLEDEBGRKAKEKNTTAETLFEINKWRLYE----- 143
 Db 1328 NALAHALQSSRHDCDLLREQYEEBQEGKAEIQRALSKANSEVAQWRTKYETDAIQTTEL 1387

QY 144 -----ELYN-----KTK----- 150
 Db 1388 EAQEKLAQRLODSEQVAVNAKASLTKTKORLQGEVEDLMVDVERANSAAALDKKQ 1447
 QY 151 -----PFQQLDAFEVKEQALLNEHGAAGQEQLNKIRDSYAKLIG-----HONLK 194
 Db 1448 RNFQKVLAEWTKCESQAELEASLESLSLTELFLKNAVEEALDQLETVKRENVKLE 1507
 QY 195 QKIHVVKLKDBNSQLKSEVSKLRQLAKK-----QSTKLOEELNKVGLGK 242
 Db 1508 QEIADLTQIAENGKTIHELEKSRQIELEKADIQIALFEAEAALEHEAKILRIQ 1563
 RESULT 14
 JX0178
 myosin heavy chain, fast skeletal muscle, adult [validated] - chicken
 N:Contains: myosin ATPase (EC 3.6.4.1)
 C:Species: Gallus gallus (chicken)
 C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 19-Apr-2002
 C:Accession: PX0050; PX0051; PX0052; JX0178; A26365; S02082; PW0009; S39081; S24351; S05
 J:Hayashida, M.; Maita, T.; Matsuda, G.
 J:Biochem. 110, 54-59, 1991
 A:Title: The primary structure of skeletal muscle myosin heavy chain: I. Sequence of the
 A:Reference number: PX0050; MUID:92041767; PMID:1939027
 A:Accession: PX0050
 A:Molecule type: protein
 A:Residues: 1-205 <HAY>
 R:Komine, Y.; Maita, T.; Matsuda, G.
 J:Biochem. 110, 60-67, 1991
 A:Title: The primary structure of skeletal muscle myosin heavy chain: II. Sequence of the
 A:Reference number: PX0051; MUID:92041768; PMID:1939028
 A:Accession: PX0051
 A:Molecule type: protein
 A:Residues: 206-636 <KOM>
 R:Maita, T.; Miyashishi, T.; Matsuzono, K.; Tanioka, Y.; Matsuda, G.
 J:Biochem. 110, 68-74, 1991
 A:Title: The primary structure of skeletal muscle myosin heavy chain: III. Sequence of the
 A:Reference number: PX0052; MUID:92041769; PMID:1939029
 A:Accession: PX0052
 A:Molecule type: protein
 A:Residues: 201-213; 632-837 <MAI>
 R:Maita, T.; Yajima, E.; Nagata, S.; Miyashishi, T.; Nakayama, S.; Matsuda, G.
 J:Biochem. 110, 75-87, 1991
 A:Title: The primary structure of skeletal muscle myosin heavy chain: IV. Sequence of the
 A:Reference number: JX0178; MUID:92041770; PMID:1939030
 A:Accession: JX0178
 A:Molecule type: protein
 A:Residues: 833-1938 <MA2>
 R:Maita, T.; Hayashida, M.; Tanioka, Y.; Komine, Y.; Matsuda, G.
 Proc. Natl. Acad. Sci. U.S.A. 84, 416-420, 1987
 A:Title: The primary structure of the myosin head.
 A:Reference number: A26365; MUID:87092420; PMID:3467365
 A:Accession: A26365
 A:Molecule type: protein
 A:Residues: 1-129, 'X', 131-139, 141-550, 'X', 552-754, 'X', 756-784, 'QL', 787-804, 806-810 <WA3>
 R:Watanabe, B.
 Biol. Chem. Hoppe-Seyler 370, 55-61, 1989
 A:Title: Amino-acid sequence of the hinge region in chicken myosin subfragment-2.
 A:Reference number: S02082; MUID:89228549; PMID:2713098
 A:Accession: S02082
 A:Molecule type: protein
 A:Residues: 1144-1270 <WAT>
 R:Yajima, E.
 Nagasaki Igakai Zasshi 65, 409-430, 1990
 A:Title: Study on tail region of skeletal
 A:Reference number: PW0009
 A:Accession: PW0009
 A:Molecule type: protein
 A:Residues: 1304-1938 <YAI>
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.
 submitted to the EMBL Data Library, August 1991
 A:Description: Analysis of the chicken fast myosin heavy chain family: Localization of i

A:Reference number: S39081
 A:Accession: S39081
 A:Molecule type: mRNA
 A:Residues: 1081-1203, 'DV', 1206-1342, 'E', 1344-1544, 'S', 1546-1795, 'HV', 1798-1829, 'S', 18
 A:Cross-references: EMBL:M74084
 R:Moore, L.A.; Arrizubieta, M.J.; Tidyman, W.E.; Herman, L.A.; Bandman, E.
 J. Mol. Biol. 225, 1143-1151, 1992
 A:Title: Analysis of the chicken fast myosin heavy chain family. Localization of isofo
 A:Reference number: S24348; MUID:92309413; PMID:1377278
 A:Accession: S24351
 A:Molecule type: mRNA
 A:Residues: 1082-1182, 'T', 1184-1203, 'DV', 1206-1342, 'E', 1344-1544, 'S', 1546-1780, 'M', 178
 A:Cross-references: EMBL:M74084
 R:Watanabe, B.
 Biol. Chem. Hoppe-Seyler 370, 1027-1034, 1989
 A:Title: Complete amino-acid sequence of subfragment-2 in adult chicken skeletal muscl
 A:Reference number: S05515; MUID:90121764; PMID:2610940
 A:Accession: S05515
 A:Molecule type: protein
 A:Residues: 842-906, 'Q', 908-1270 <WA3>
 R:Watanabe, B.
 Biol. Chem. Hoppe-Seyler 370, 549-558, 1989
 A:Title: Amino-acid sequence of the short subfragment-2 in adult chicken skeletal musc
 A:Reference number: S04501; MUID:89374803; PMID:2775482
 A:Accession: S04501
 A:Molecule type: protein
 A:Residues: 852-906, 'Q', 908-1108 <WA2>
 R:Matsuda, G.; Maita, T.; Miyashishi, T.; Hayashida, M.
 J. Protein Chem. 6, 33-46, 1987
 A:Title: Structure and function of muscle myosin.
 A:Reference number: A60877
 A:Accession: A60877
 A:Molecule type: protein
 A:Residues: 1-139, 141-205 <MA1>
 R:Gulick, J.; Kropp, K.; Robbins, J.
 J. Biol. Chem. 260, 14513-14520, 1985
 A:Title: The structure of two fast-white myosin heavy chain promoters. A comparative s
 A:Reference number: A92507; MUID:86039556; PMID:2997212
 A:Accession: A24124
 A:Molecule type: DNA
 A:Residues: 'M', 1-168 <GUL>
 A:Cross-references: GB:M13512; GB:M12083; GB:M13510; NID:G212363; PIDN:AAA48966.1; PID
 R:Kropp, K.; Gulick, J.; Robbins, J.
 J. Biol. Chem. 261, 6613-6618, 1986
 A:Title: A canonical sequence organization at the 5'-end of the myosin heavy chain gen
 A:Reference number: A92587; MUID:86196091; PMID:3009465
 A:Accession: C25217
 A:Molecule type: DNA
 A:Residues: 'M', 1-56, 'T', 58-76, 'I', 78-168 <KRO>
 A:Cross-references: GB:M13515; GB:M13511; NID:G212373; PIDN:AAA48971.1; PID:G555468
 C:Comment: This is a fragment of the globular head.
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: acetylated amino end; actin binding; ATP; coiled coil; hydrolase; methylat
 F:1-1938/Product: myosin heavy chain #status experimental <MAT>
 F:89-768/Domain: myosin motor domain homology <MYOT>
 F:179-186/Region: nucleotide-binding motif A (P-loop)
 F:1550-587/Region: actin binding #status predicted
 F:657-679/Region: actin binding #status predicted
 F:841-1938/Domain: actin binding #status predicted
 F:841-1289/Region: S2
 F:852-1108/Domain: short subfragment 2 <SUB2>
 F:1290-1938/Region: light meromyosin
 F:1/Modified site: acetylated amino end (Ala) #status experimental
 F:130,551/Modified site: N6-methyllysine (lys) #status experimental
 F:185/Binding site: ATP (lys) #status predicted
 F:697, 707/Active site: Cys #status predicted
 F:755/Modified site: 3'-methylhistidine (His) #status experimental
 Query Match 14.9%; Score 177.5; DB 1; Length 1938;
 Best Local Similarity 20.5%; Pred. No. 0.12;
 Matches 73; Conservative 55; Mismatches 107; Indels 121; Gaps 10;

Db 1257 DVSELTKEBEEQORLINDLTAQARLQTEAGEYSRQDLKDALYSQLSRSKQASTQOIE 1316
 QY 98 DLQNLQKOEED-----FRKQLEDEGRKAKEKNTTAEITTEINKWRL 140
 Db 1317 ELKHQLEETAKNALAHALOSSRRHDCDLLREQYEEQEGKALQRLSKANSEVAQWRT 1376
 QY 141 LY-----BELYNKTKPFOIQD-----APEVEKQALLNE----- 169
 Db 1377 KYETDAIORTLEBAKKLQRLQREABEHVAVNAKASLEKTKQRLONEVEDMLDVE 1436
 QY 170 HGAQEQOLKIRDSYAKLLGHONLKQ-----KIKHVVK----- 202
 Db 1437 RSNAAACALDKQRFNVLSWKQYETQAELEASQKESLSLSTELFKVKNVVEBSLD 1496
 QY 203 ----LKDNSQLKSVSKLRQLAK-----KKQ-----SETKLOE 233
 Db 1497 QLETLRENTQNLQOISDLTEQIABGGKQIHELEKIKQVQBEKCEIQAALEAEASLEH 1556
 QY 234 ELNKVLGK 242
 Db 1557 BEGKILRIQ 1565

RESULT 17
 A29320
 myosin heavy chain, fast skeletal muscle, embryonic [similarity] - chicken
 N;Alternate names: myosin heavy chain, EFWI
 C;Species: Gallus gallus (chicken)
 C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 02-Feb-2001
 C;Accession: A29320; B24124; A02990
 R;Molina, M.I.; Kropp, K.E.; Gulick, J.; Robbins, J.
 J. Biol. Chem. 262, 6478-6488, 1987
 A;Title: The sequence of an embryonic myosin heavy chain gene and isolation of its corre
 A;Reference number: A29320; MUID:87194881; PMID:3571266
 A;Accession: A29320
 A;Molecule type: DNA; mRNA
 A;Residues: 1-1940 <MOL>
 A;Note: the sequence translated from the mRNA differs from that of the DNA in having 379
 R;Gulick, J.; Kropp, K.; Robbins, J.
 J. Biol. Chem. 260, 14513-14520, 1985
 A;Title: The structure of two fast-white myosin heavy chain promoters. A comparative stu
 A;Reference number: A92507; MUID:86033956; PMID:2997212
 A;Accession: B24124
 A;Molecule type: DNA
 A;Residues: 1-168 <GUI>
 R;Kavinsky, C.U.; Umeda, P.K.; Sinha, A.M.; Elzinga, M.; Tong, S.W.; Zak, R.; Jakovcic,
 J. Biol. Chem. 258, 5196-5205, 1983
 A;Title: Cloned mRNA sequences for two types of embryonic myosin heavy chains from chick
 A;Reference number: A02990; MUID:83161144; PMID:6833296
 A;Accession: A02990
 A;Molecule type: mRNA
 A;Residues: 1502-1546, 'A', 1548-1912, 'G', 1914, 'T', 1916-1940 <XAV>
 A;Cross-references: GB:V00430; GB:J00892; NID:G63613; PIDN:CAA23712.1; PID:g808972
 A;Experimental source: clone 251
 C;Comment: The entire sequence of this fragment corresponds to the light meromyosin port
 ides (heptads I-IV), that are characteristic of coiled coils.
 C;Comment: Association of the alpha-helical regions of two myosin heavy chains to form a
 C;Superfamily: myosin heavy chain; myosin motor domain homology
 C;Keywords: actin binding; ATP; coiled coil; methylated amino acid; muscle; muscle contr
 F;89-771/Domain: myosin motor domain homology <MMOT>
 F;179-186/Region: nucleotide-binding motif A (P-loop)
 F;130/Modified site: N6,N6,N6-trimethyllysine (lys) #status predicted

Query Match 14.8%; Score 176.5; DB 2; Length 1940;
 Best Local Similarity 20.2%; Pred. No. 0.13;
 Matches 72; Conservative 53; Mismatches 110; Indels 121; Gaps 9;

QY 8 VQSLSDVTAQESYKALTASIEDIKLENSLOEKAAKAGNAEDVQHQILATSSNOEY 67
 Db 1212 IDNLQVRQKLEKESELKNEIDDIASNNVESYKAKANLEKWCRLSDQLSEIKTEKEEQ 1271
 QY 68 VRMLDLQTKGALKETEIKET-----VSFLOKITDLQNLQKQ----- 106

Db 1272 QRTINDISAQXARLQTSGEYSRQVEBKDALISQLSRKQAFQIQIIELEKRLHEEIEKAK 1331
 QY 107 -----EED-----FRKQLEDEGRKAKEKNTTAEITTEINKWRLY-----EEL 145
 Db 1332 KCPAHALQSARHDCDLLREQYEEQEGKALQRLSKANSEVAQWRTKYETDAIORTTEL 1391
 QY 146 YNKTQPIQD-----APEVEKQALLNE-----HGAQEQOLNKIR 181
 Db 1392 EAKKKLAQRLQDAEBHVEAVNSKASLEKTKQRLONEVEDLMIDVERSNAACALDKKQ 1451
 QY 182 DSYAKLLGHONLK-----KKQ-----OKIKHVVKLDKENSOLK 211
 Db 1452 KNFKKILSEWKQYETQAELEASQKESRSLSLSTELFKMKNAYEESLDHLETKENKQ 1511
 QY 212 SEVSKLRQLAK-----KKQ-----SETKLOEINKVLGK 242
 Db 1512 QEISDLTEQIABGGKQIHELEKIKQVQBEKCEIQAALEAEASLEHEGKILRIQ 1567

RESULT 18
 A23767
 myosin heavy chain, fast skeletal muscle - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 20-Jun-2000
 C;Accession: A23767
 R;Saez, L.; Leinwand, L.A.
 Nucleic Acids Res. 14, 2951-2969, 1986
 A;Title: Characterization of diverse forms of myosin heavy chain expressed in adult hu
 A;Reference number: A93616; MUID:86176778; PMID:2421254
 A;Accession: A23767
 A;Molecule type: mRNA
 A;Residues: 1-876 <SAB>
 A;Cross-references: GB:X03740; NID:g34839; PIDN:CAA27380.1; PID:g1335218
 C;Superfamily: myosin heavy chain; myosin motor domain homology
 C;Keywords: actin binding; ATP; coiled coil; muscle; skeletal muscle

Query Match 14.7%; Score 175.5; DB 2; Length 876;
 Best Local Similarity 21.9%; Pred. No. 0.064;
 Matches 79; Conservative 50; Mismatches 102; Indels 129; Gaps 11;

QY 8 VQSLSDVTAQESYKALTASIEDIKLENSLOEKAAKAGNAEDVQHQILATSS 63
 Db 148 IDNLQVRQKLEKESELKNEIDDIASNNVESYKAKANLEKWCRLSDQLSEIKTEK 203
 QY 64 NOEYVRMLDLQTKSA-----LKTEIKETVSPFLQKITDLQNLQKQ 106
 Db 204 EEEQRLINDLTAQARLQNOVEYSRQDLKDTLETQLSGKQAFQIQIIELEKRLHEE 263
 QY 107 -----EED-----FRKQLEDEGRKAKEKNTTAEITTEINKWRLY----- 142
 Db 264 IKAKSALAHALQSSRHDCDLLREQYEEQEGKALQRLSKANSEVAQWRTKYETDAIQR 323
 QY 143 EELYNKTKPFOIQD-----APEVEKQALLNE-----HGAQEQOL 177
 Db 324 TEELERAKKLQRLQDAEBHVEAVNAKASLEKTKQRLONEVEDLMIDVERTNAACAL 383
 QY 178 NKIRDSYAKLLGHONLKQIKHV-----KLKDN 207
 Db 384 DKQTNFKILAEWKQCEETHAVLESFQKESLSLSTELFKMKNAYEESLDQETLKR 443
 QY 208 SOLKSEYKLRQLAK-----KKQ-----SETKLOEINKVLGK 242
 Db 444 KNLQOISDLTEQIABGGKQIHELEKIKQVQBEKCEIQAALEAEASLEHEGKILRIQ 503

RESULT 19
 S49461
 synaptonemal complex protein 1 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
 C;Accession: S49461; S59599
 R;Julien, S.; Luc, M.; Francois, C.
 submitted to the EMBL Data Library, October 1994

A;Description: Cloning and sequencing of the murine SCPI cDNA.

A;Reference number: S49461

A;Accession: S49461

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-993 <JUL>

A;Cross-references: EMBL:Z38118; NID:G1360015; PIDN:CAA86262.1; PID:G558603

R;Sage, J.; Martin, L.; Cuzin, F.; Rassoulzadegan, M.

Biochim. Biophys. Acta 1263, 258-260, 1995

A;Title: cDNA sequence of the murine synaptonemal complex protein 1 (SCPI).

A;Reference number: S59599; MUID:96004899; PMID:7548215

A;Accession: S59599

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-993 <SAG>

A;Cross-references: EMBL:Z38118; NID:G1360015; PIDN:CAA86262.1; PID:G558603

Query Match 14.6%; Score 175; DB 2; Length 993;

Best Local Similarity 21.3%; Pred. No. 0.077;

Matches 73; Conservative 68; Mismatches 85; Indels 116; Gaps 14;

QY 11 LEDVTAQFESYKALTASIEDLKLENSLSQEAQAKGNAAE-----51

DB 416 LEEMT-KFKNKEVELEELKNILAEQKLLDEKKQVKEKLAELOEKEQELTFLLETRKE 474

QY 52 --DVQHOILATESNQEYVRMLDLDT---KALKETEIKEITVS-----FLQKI 96

DB 475 VHDLOEQVTVTKTSEGHYKQVEEMKTELEKLNKT---ELTASCDMLLENKFKVQEA 531

QY 97 TDLQNLQKQEBD-----PRKQLE-----DEGRKA 122

DB 532 SMALELKHQSDIINCKQOERLLKQIENBEKEMHLRDELESVRKEFIQGDVYCKL 591

QY 123 EKENTTA-----ELTBEINKRWLL-----YEELYNKTKPPQ-----153

DB 592 DKSEENARSIECVLKKQKQKILSKONKLLQVENKSKNIEELHQENKTLKKSSAEI 651

QY 154 IOLDAFEVQKALLNEHGAQAQOLNIRDSYAKLLGHQNLQK--KIKHVVKLK---DENS 208

DB 652 KQLNAYEIKVSKLELESTKQRFERMTNNYQKEIENKKIISGKLLGEVEKAKATVDEAV 711

QY 209 QLKSVSKLRQ-----LAKKQKQSETKLOEELNKVLGI 241

DB 712 KLOKSID-LRCQKHAEMVMALEKHKHQYDKIVEERDSBLG 752

RESULT 20

A60608

paramyosin - fluke (Schistosoma mansoni) (fragment)

C;Species: Schistosoma mansoni

C;Date: 17-Apr-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jan-2000

C;Accession: A60608; A25993

R;Grossman, Z.; Ram, D.; Markovics, A.; Tarrab-Hazdai, R.; Lantner, F.; Ziv, E.; Schecht

Exp. Parasitol. 70, 62-71, 1990

A;Title: Schistosoma mansoni: stage-specific expression of muscle-specific genes.

A;Reference number: A60608; MUID:90108200; PMID:2295327

A;Accession: A60608

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-439 <GRO>

A;Note: authors also sequenced a portion of the mRNA described in reference A25993 and

R;Lanar, D.E.; Pearce, E.J.; James, S.L.; Sher, A.

Science 234, 593-596, 1986

A;Title: Identification of paramyosin as schistosoma antigen recognized by intradermally

A;Reference number: A25993; MUID:87018840; PMID:3094144

A;Accession: A25993

A;Molecule type: mRNA

A;Residues: 1-120, 'GSSQRIRKLL', 130-334, 'L', 336, 'I', 338-424, 'R', 426-439 <LAN>

A;Cross-references: GB:M14163; NID:G161054; PID:G161055

A;Note: this sequence has been revised in reference A60608

C;Superfamily: myosin heavy chain; myosin motor domain homology

C;Keywords: coiled coil; muscle

Query Match 14.4%; Score 172.5; DB 2; Length 439;
Best Local Similarity 22.6%; Pred. No. 0.044;
Matches 77; Conservative 50; Mismatches 96; Indels 117; Gaps 12;

QY 1 QEKYDSMVQSLSDVTAQFESYKALT-----ASEIEDLKLENSLSQEAQAK 45

DB 4 KRKTMITELED-TAERERLKAVSLKLEKTKLTKLTKLQSEIESLSLENSLEIRAKA 62

QY 46 AGKNAEDVQHOI-----LATESNQEYVRMLDLDTQKSAKETE-----84

DB 63 AESLASDLQRRVDELTTIEVNTLTSONQSELENRLKSLVNDLTDKNLLERENQMDQ 122

QY 85 IKEITVSPFL---QKITDLO-----NQL 103

DB 123 VKELSSURDANRLTDLALRSQLEAFRDNLASALHDAEALHDMQOYQASQALNHL 182

QY 104 K-OOEEDPRKOLEDEGRKAENKTAEITBEINKRWLLYE-ELYNKTKPQIOLDAFEV 161

DB 183 KSEMQRLREDELESURKSTTTTIELTWTIEMEVKYKSELRLKRYESNIADLEI 242

QY 162 EKQALLNEHGAQAQOLNIRDSYAKLL-GHONLKQKIKHVVKLKDENS-----208

DB 243 -----QLDTANKANANLKENKNSQVKDLETFLEDEERRLEAAENNLQI 288

QY 209 -----QLKSVSKLRQOLAK-----KKQSETKLOEELNKV 238

DB 289 TEHKLQLANEIEIRSTLENLERLRKHAETELBEAQSRV 328

RESULT 21

B30395

purine NTPase [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: B30395

R;She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Cha

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,

arrest, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-864 <KUR>

A;Cross-references: GB:AE006641; NID:G13815551; PIDN:AAK42417.1; GSPDB:GN00155

C;Genetics:

A;Gene: SSO2249

Query Match 14.4%; Score 172.5; DB 2; Length 864;

Best Local Similarity 25.3%; Pred. No. 0.089;

Matches 68; Conservative 51; Mismatches 99; Indels 51; Gaps 10;

QY 2 EKYDSMVQSLSDVTAQFESYKALTASETE---DLKLENSLSQ-EKAAKAGNAE-----51

DB 272 EKEINLENLNIKLFKEYEVLAQSHTEMANVINLEKEIEEYKAIIRKELEPKVLK 331

QY 52 --DVQHOILATESNQEYVRMLDLDTQKSAKETEIKEITVSFLQKITDLQNLQKQEBD 109

DB 332 YKELERKLEELQPKYQYVLKSLDLSKLNLERLEKD--ASELNDIDKVNLSLQKVEE 389

QY 110 FRKQLEDEGRKAENKTAEITBEINK-----WELLYEE-----LYNKTKPQ 153

DB 390 TRKQNLNRAQLAKVESLISEKNEIINNISQVEGTCVCGRPLDEEHKQKIKRAKSYI 449

QY 154 IOLDAFEVQKALLNEHGAQAQOLNIRDSYAKL---LGHONLKQKIKHVVKLKDENSQ 209

DB 450 LQL-----ELNKLELEELKKITNELNKIEFRRLSNVNVASVDNVRQLK---LNKEETEN 503

QY 210 LKSEVSKLRQOLAKKKQKQSETKLOEELNKV 238

DB 504 LHSEIESLK-----NIDEIKKI 521

```

Db      97  QRTITLDTLTTQAAKLTQENSELSRQLBEKLNWLRGKLSLTQQLEDLKKRLQEEAKARN 156
Qy      107  -----BED---FRQLDEEGRKAEKENTTAELTEEINKWRLLY-----EELY 146
Db      157  ALAHALQSAQHDCDLLRQYEEEMAEKALQALSKANSEVAQWRTKYETDAIQRTEELE 216
Qy      147  NKTFFQIQL-----DAFEVKKQALLNE-----HGAAQQLNKIRD 182
Db      217  EAKKLAQRLOEABEAVNAKCSSLEKTKHRLQNEVEDLMDADVERSNAAAAALDKQR 276
Qy      183  SYAKLIGH-----QNLKQKHVVUKDSENSQLKS 212
Db      277  NFDKILSEWKQKFESQTELEASQKEARSLSSTELFKLKKSYYELMEHLSTFFRENKNLQE 336
Qy      213  EVS-----KLRCLQAKKK-----QSETKLQELNKNVL 239
Db      337  EISGLTBLGVQOKSIHLEKVRKQLDREKLELEAALAEAEASLHEEGKIL 388

RESULT 24
A59234
slow myosin heavy chain 3 - quail
C:Species: Coturnix coturnix
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
C:Accession: A59234
R:Nikowitz Jr., W.; Wang, G.F.; Feldman, J.L.; Miller, J.B.; Wade, R.; Nelson, L.; Sto
J. Biol. Chem. 271, 17047-17056, 1996
A:Title: Isolation and characterization of an avian slow myosin heavy chain gene expre
A:Reference number: A59234; MUID:96291845; PMID:8663323
A:Accession: A59234
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1931 <NIK>
A:Cross-references: GB:U53862; NID:g1289513; PID: AAC59912.1; PID:g1289514
C:Superfamily: myosin heavy chain; myosin motor domain homology
F:81-761/Domain: myosin motor domain homology <MMC>

Query Match      14.1%; Score 169; DB 2; Length 1931;
Best Local Similarity 21.9%; Pred. No. 0.31;
Matches 75; Conservative 51; Mismatches 112; Indels 104; Gaps 10;

Qy      1  QEKYDSMVQSLVEDVTAFESYKALTASIEDLKLK---NSSLOEKAAKAGKVAE----DV 53
Db      1188  RKGHADVAELSEQLDNMQRVQKLEKEKSELKLELDVSSNNQELIKAKANLEKMCNST 1247
Qy      54  QHQLATESSNOEYVRMLDLDQTSALKETEKIT-----VSFLOKI 96
Db      1248  EDQNNHKNKLESQRTVDTLSTORAKLTQENSELSROLEKEAFINQMRGLTITQOL 1307
Qy      97  TDLQNLKQQ-----BED---FRQLDEEGRKAEKENTTAELTEEINKWR 139
Db      1308  EDLKRQLEEEAKAKNALAHALQSAQHDCDLLRQYEEEMAEKALQALSKANSEVAQWR 1367
Qy      140  LLY-----HELYNTKFFQIQL-----DAFEVKKQALLNE----- 169
Db      1368  TKYETDAIQRTELEAEAKKLAQRLOEABEAVNAKCSSLEKTKHRLQNEVEDLMDADV 1427
Qy      170  --HGAAQQLNKIRDSYAKLIGHQNLK-----QKI 197
Db      1428  ERSNAAAALDKQRNPKDILSEWKQKFESQTELEASQKEARSLSSTELFKLNAYEESL 1487
Qy      198  KHVVKDKENSQKSEYSKLRCQAKKQSKETKLQELNKNVL 239
Db      1488  EHLETFRENKNLQEEISDLTEQLGASQKSIHEL-EKVRKQL 1528

RESULT 25
T18278
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000
C:Accession: T18278; T30579
R:Hammer III, J.A.; Jung, G.

```

J. Biol. Chem. 271, 7120-7127, 1996
A;Title: The sequence of the dictyostelium myo J heavy chain gene predicts a novel, dime
A;Reference number: Z18854; MUID:96215148; PMID:8636147
A;Accession: T18278
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2245 <HAM>
A;Cross-references: EMBL:U42409; NID:g1150765; PID:g1150766; PIDN:AAA85186.1
R;Titus, M.A.; Kuspa, A.; Loomis, W.F.
Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994
A;Title: Discovery of myosin genes by physical mapping in Dictyostelium.
A;Reference number: Z20873; MUID:95023328; PMID:7937787
A;Accession: T30579
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-190, 'F', 192-283, 'T', 285-290, 'R', 292-331, 'IE', 'WM', 337-338, 'LK', 342, 'YRMS',
A;Cross-references: EMBL:L35322; NID:g1039360; PID:g1039361; PIDN:AAA79858.1
C;Genetics:
A;Gene: myoJ
A;Introns: 257/3; 307/3
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: ATP; P-loop
F;84-809/Domain: myosin motor domain homology <MMO>

Query Match 14.0%; Score 167; DB 2; Length 2245;
Best Local Similarity 22.1%; Pred. No. 0.46;
Matches 56; Conservative 60; Mismatches 113; Indels 24; Gaps 6;
QY 2 EKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLOEKAAGKNAEDVQHOILATE 61
DB 1118 QOLDENKBEFDELSDERTDNTNQLQKANKSTLEEDYFSLGIDNRLERQVLEIR 1177
QY 62 SSNOYVRMLD-----LQTKALKETEIKEITVSVFLQKITDLQNLQKQEEEDPRKQ 113
DB 1178 DENQ-LIKERLDSLGQSSQFQSGAALKQEQLEQVQEQEQLIKLSSEKLGSEBEAKKQ 1236
QY 114 -----LEDEGRKAERKNTTAETAEINRWLYEELYNKTKFQOLDFAFEVKQAL 166
DB 1237 INQLELTDHRSKLIQIQLTQEQNEKTKLGLKLEEQDEKQLQEQLEIRIKOSKQSV 1296
QY 167 LNEHGAQQLNKIRDSYAKL--LGHQNLKOKIKHWVKLDKENSOLKSEVSKLRQALAK 223
DB 1297 EDEKNSLTQLTTFESTQVSTNVSHQ--KEKIL---TTLKSTIELNKSIGKLAQKQ 1351
QY 224 KQSETKLQEEELN 236
DB 1352 KQDEIRKIQFELN 1364.

RESULT 26
S18199
myosin heavy chain - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 02-Feb-2001
C;Accession: S18199
R;Stewart, A.F.R.; Camoretti-Mercado, B.; Perlman, D.; Gupta, M.; Jakovcic, S.; Zak, R.
J. Mol. Evol. 33, 357-366, 1991
A;Title: Structural and phylogenetic analysis of the chicken ventricular myosin heavy ch
A;Reference number: S18199; MUID:92130260; PMID:1774788
A;Accession: S18199
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1039 <STE>
A;Cross-references: EMBL:X59552; NID:g62995; PIDN:CAA42130.1; PID:g62996
A;Note: In the authors' translation 45-Lys is shown after residue 40, and, consequently,
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: ATP; nucleotide binding; P-loop
F;402-409/Region: nucleotide-binding motif A (P-loop)

Query Match 13.9%; Score 166.5; DB 2; Length 1039;
Best Local Similarity 20.9%; Pred. No. 0.22;
Matches 71; Conservative 55; Mismatches 94; Indels 119; Gaps 11;

QY 8 VQSLIEDVTAQFESYKALTASEIEDLKLENSLOEKAAGKNAEDV-----QHOILATESS 63
DB 310 LDNLQVRQKLEKEKSELKMEVDL-----TANMEQTVKGNAAEKLWGTVYEDHLNETKTK 365
QY 64 NOEYVRMLDLQTKSA-----LKETIKETIV-----SFLQKITDLQNLQKQ 106
DB 366 LDEWTRLMNDLITQTKLKQSENGEFVQLEEKESLISQLSRGKTSFTQOIEELARQLEEE 425
QY 107 EED-----FRQLEDEGRKAERKNTTAETAEINRWLYEELYNKTKFQOLDFAFEVKQAL 142
DB 426 TSKNNAHALQAARHDCDLLREQYEEQEAQALQALSKGNAEVAQWRTKYETDAIQ 485
QY 143 -EELYNKTKFQIQOL-----DAFEVKQALINEH-----GAQEQOL 177
DB 486 TEELDEAKKLLAKLQAEAEAEANAKCSLEKAKHRLQNEQEDMMIDLEKANSAAASL 545
QY 178 NKIRDSYAKLIGHONLQK-----NDWQKYEESQAELEASQEAESLSSTELFKLNAVEETLDHLETLKR 603
DB 546 DKQGRGDKII--NDWQKYEESQAELEASQEAESLSSTELFKLNAVEETLDHLETLKR 603
QY 206 ENSOLKSEVSKLRQALAK-----KKQSETKLQ 233
DB 604 ENKNLQBEISDLTNQISEGNKNLHEIEKVKKQVEQEKSE 642

RESULT 27
I53799
CG1 protein - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C;Accession: I53799
R;Print, C.G.; Leung, E.; Harrison, J.E.; Watson, J.D.; Krissansen, G.W.
Gene 144, 221-228, 1994
A;Title: Cloning of a gene encoding a human leukocyte protein characterised by extensi
A;Reference number: I53799; MUID:94314220; PMID:8039706
A;Accession: I53799
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1300 <RES>
A;Cross-references: GB:L25616; NID:g409465; PIDN:AA65853.1; PID:g409466
C;Genetics:
A;Gene: GDB.CG1
A;Cross-references: GDB:450341

Query Match 13.9%; Score 166; DB 2; Length 1300;
Best Local Similarity 24.8%; Pred. No. 0.29;
Matches 76; Conservative 50; Mismatches 98; Indels 83; Gaps 14;

QY 1 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLOEKA--AKAGKNAEDVQHOIL 58
DB 452 KQYARLVNLETKTGK-----LQEEVQKKNAAEQAAATQKVLQEAERWEEVQSYI- 504
QY 59 ATESSNOEYVRMLDLQTKSALKETEIKEITVSVFLQKITD-----LQNLQKQ----- 105
DB 505 --RKETAHEAAQQDLQSKFVAKENEVQ-----SLHSLKLTDLVSKQQLQRLMQLMSESEQ 558
QY 106 -----QBEDFRKQ-----LEDEGRKAERKNTTAETAEINRWLYEELYNKTKFQOLDFAFEVKQAL 153
DB 559 KRVNKEESLQMQVDILQEQEALKAQIQPFHSAQAATSA-SVLABELHFWIAEKDKQIK 617
QY 154 LQDAF-----EVEK-QALLNEHGAQEQNLK-----I 180
DB 618 QTESLASERDLRTSKBEELKDQINMFLKAEVQKQLQALANQAAAHELEKQOQSYV 677
QY 181 RDSYAKLIGHQ---NLKQIKHWVKLDKENSOLKSEVSKLRQALAK-----KQSETKL 231
DB 678 KDKIRLLEEQQLQHEISNKEEFKILNDQNKALKSEVQKQLTLVSEQPKNDVWEQMEKCI 737
QY 232 QEELNKV 238
DB 738 QEKDEKL 744

Search completed: December 16, 2003, 06:08:37
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:07:28 ; Search time 18 Seconds
(without alignments)

632.248 Million cell updates/sec

Title: US-09-978-309A-74

Perfect score: 1195

Sequence: 1 QEYDQNMVQSLDVTQAQFES.....KKXQSETKLQELNKVLGK 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1193	99.8	724	1	HMHR_HUMAN
2	917	76.4	794	1	HMHR_MOUSE
3	893	75.1	498	1	HMHR_RAT
4	191	16.0	976	1	SCP1_HUMAN
5	189	15.8	1690	1	C190_DROME
6	185.5	15.5	1790	1	USO1_YEAST
7	182.5	15.3	1938	1	MYH4_RABIT
8	181.5	15.2	1940	1	MYH3_RAT
9	180.5	15.1	1941	1	MYH2_HUMAN
10	179.5	15.0	1938	1	MYSS_CHICK
11	179.5	15.0	1939	1	MYH4_HUMAN
12	179	15.0	697	1	MEP1_LYCES
13	178.5	14.9	1939	1	MYH1_HUMAN
14	178.5	14.9	1940	1	MYH3_HUMAN
15	176.5	14.8	845	1	SCP1_MESAU
16	176.5	14.8	1937	1	MYH8_HUMAN
17	176.5	14.8	1940	1	MYH3_CHICK
18	175	14.6	993	1	SCP1_MOUSE
19	172.5	14.4	864	1	RA50_SULSO
20	172.5	14.4	866	1	MYSP_SCHMA
21	171.5	14.4	1935	1	MYSS_CYPCA
22	171	14.3	886	1	RA50_SULAC
23	169	14.1	2230	1	GOG4_HUMAN
24	167	14.0	2245	1	MYSJ_DICDI
25	166.5	13.9	1084	1	MYSS_RABIT
26	166.5	13.9	1102	1	MYSS_CHICK
27	165.5	13.8	879	1	RA50_SULTO
28	165	13.8	1360	1	CING_XENLA
29	162.5	13.6	1938	1	MYH6_MOUSE
30	162.5	13.6	1939	1	MYH6_HUMAN
31	162	13.6	1935	1	MYH7_PIG
32	161.5	13.5	1938	1	MYH6_RAT
33	161.5	13.5	1939	1	MYH6_MESAU

ALIGNMENTS

RESULT 1

HMHR_HUMAN

ID HMHR_HUMAN STANDARD; PRT; 724 AA.

AC QY5330; Q92767;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hyaluronan mediated motility receptor (Intracellular hyalurononic acid binding protein) (Receptor for hyaluronan-mediated motility) (CD168 antigen).

DE antigein.

GN HMHR OR IHABP OR RHAMM.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RC SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.

RA TISSUE=Breast carcinoma;

RA MEDLINE=98264864; PubMed=9601098;

RA Asmann V., Marshall J.F., Fieber C., Hofmann M., Hart I.R.;

RT "The human hyaluronan receptor RHAMM is expressed as an intracellular protein in breast cancer cells.";

RL J. Cell Sci. 111:1685-1694(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Breast;

RC MEDLINE=97045829; PubMed=8890751;

RA Wang C., Entwistle J., Hou G., Li Q., Turley E.A.;

RT "The characterization of a human RHAMM cDNA: conservation of the hyaluronan-binding domains.";

RL Gene 174:299-306(1996).

CC -! FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO

CC HMHR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE

CC FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR

CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING

CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.

CC -! SUBUNIT: SUBUNIT OF THE HARC COMPLEX.

CC -! SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC (BY

CC SIMILARITY).

CC -! ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=A;

CC IsoId=QY5330-1; Sequence=Displayed;

CC Name=B;

CC IsoId=QY5330-2; Sequence=VSP_004286;

CC -! TISSUE SPECIFICITY: EXPRESSED IN BREAST CANCER CELL LINES AND IN

CC NORMAL BREAST TISSUE.

CC -! DATABASE: NAME=PROW; NOTE=PROW 2:76-84(2001);

CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/802868666_g.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use. by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF032862; AAC32548.1; --
 DR EMBL; U29343; AAC52049.1; --
 DR Genew; HGNC:5012; HMMR.
 DR MIM; 600936; --
 DR GO; GO:0007048; P:oncogenesis; TAS.
 KW Hyaluronic acid; Alternative splicing; Repeat; Glycoprotein; Antigen.
 FT DOMAIN 635 645 HYALURONIC ACID-BINDING (POTENTIAL).
 FT DOMAIN 657 666 HYALURONIC ACID-BINDING (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 477 477 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 588 588 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARSPLIC 75 90 Missing (in isoform B).
 FT CONFLICT 75 75 /FTIGVSP_004286.
 FT CONFLICT 103 103 K -> KK (IN REF. 2).
 FT CONFLICT 277 277 S -> R (IN REF. 2).
 FT CONFLICT 298 298 E -> D (IN REF. 2).
 FT CONFLICT 322 322 K -> T (IN REF. 2).
 FT CONFLICT 330 332 K -> E (IN REF. 2).
 FT CONFLICT 330 332 QER -> REH (IN REF. 2).
 SQ SEQUENCE 724 AA; 84031 MW; EA68AD6D2A626926 CRC64;
 Query Match 99.8%; Score 1193; DB 1; Length 724;
 Best Local Similarity 99.6%; Pred. No. 2.6e-53;
 Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QKYDSWQSLSDVTAQFESYKALTASEIEDLKLENSLSQEAAGKNAEDVQHQILAT 60
 DB 442 QKYDSWQSLSDVTAQFESYKALTASEIEDLKLENSLSQEAAGKNAEDVQHQILAT 501
 QY 61 ESSNOEYVRMLDLQTSALKETEITVSTFLOKITDLONLKOQEEEDFRKOLEDEGR 120
 DB 502 ESSNOEYVRMLDLQTSALKETEITVSTFLOKITDLONLKOQEEEDFRKOLEDEGR 561
 QY 121 KAEKNTAELTEENKWLLELYNKKTPQIOLDAFEVEKQALLNEHGAQQLNKI 180
 DB 562 KAEKNTAELTEENKWLLELYNKKTPQIOLDAFEVEKQALLNEHGAQQLNKI 621
 QY 181 RDSYAKLGHQNLKQIKHVKWKDENSQKSEVSKLQCLAKKQKSETKLQEEELNKVLG 240
 DB 622 RDSYAKLGHQNLKQIKHVKWKDENSQKSEVSKLQCLAKKQKSETKLQEEELNKVLG 681
 QY 241 IK 242
 DB 682 IK 683
 RESULT 2
 HMMR MOUSE
 ID HMMR MOUSE STANDARD; PRT; 794 AA.
 AC Q00547;
 DT 01-OCT-1996 (Rel. 34; Created)
 DT 30-MAY-2000 (Rel. 39; Last sequence update)
 DT 28-FEB-2003 (Rel. 41; Last annotation update)
 DE Hyaluronan mediated motility receptor (intracellular hyaluronic acid
 DE binding protein) (Receptor for hyaluronan-mediated motility).
 GN HMMR OR IHABP OR RHAMM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
 RP TISSUE=Lung;
 RX MEDLINE=98264863; PubMed=9601097;
 RA Hofmann M., Fieber C., Assmann V., Goettlicher M., Sleeman J.,

RA Plug R., Howells N., von Stein O., Ponta H., Herrlich P.;
 RT "Identification of IHABP, a 95 kDa intracellular hyaluronate binding
 RT protein.";
 RL J. Cell Sci. 111:1673-1684 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zhao Y., Zhang S., Turley E.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-183 FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=129/SV;
 RX MEDLINE=99107769; PubMed=9889313;
 RA Fisher C., Plug R., Sleeman J., Ball P., Ponta H., Hofmann M.;
 RT "Characterization of the murine gene encoding the intracellular
 RT hyaluronan receptor IHABP.";
 RL Gene 226:41-50 (1999).
 RN [4]
 RP SEQUENCE OF 164-794 FROM N.A., AND ALTERNATIVE SPLICING.
 RC STRAIN=BALB/C; TISSUE=Fibroblast;
 RX MEDLINE=96011639; PubMed=7590272;
 RA Entwistle J., Zhang S., Yang B., Wong C., Li Q., Hall C.L., A J.,
 RA Mowat M., Greenberg A.H., Turley E.A.;
 RT "Characterization of the murine gene encoding the hyaluronan receptor
 RT RHAMM.";
 RL Gene 163:233-238 (1995).
 RN [5]
 RP SEQUENCE OF 318-794 FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=BALB/C;
 RX MEDLINE=9229690; PubMed=1376732;
 RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
 RA Cripps V., Austen L., Nance D.M., Turley E.A.;
 RT "Molecular cloning of a novel hyaluronan receptor that mediates tumor
 RT cell motility.";
 RL J. Cell Biol. 117:1343-1350 (1992).
 RN [6]
 RP CHARACTERIZATION.
 RX MEDLINE=94308286; PubMed=7518470;
 RA Hall C.L., Wang C., Lange L.A., Turley E.A.;
 RT "Hyaluronan and the hyaluronan receptor RHAMM promote focal adhesion
 RT turnover and transient tyrosine kinase activity.";
 RL J. Cell Biol. 126:575-588 (1994).
 RN [7]
 RP ERK REGULATION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=9822222; PubMed=9556628;
 RA Zhang S., Chang M.C., Zylka D., Turley S., Harrison R., Turley E.A.;
 RT "The hyaluronan receptor RHAMM regulates extracellular-regulated
 RT kinase.";
 RL J. Biol. Chem. 273:11342-11348 (1998).
 RN [8]
 RP REVIEW.
 RX MEDLINE=99059494; PubMed=9845361;
 RA Hofmann M., Assmann V., Fieber C., Sleeman J.P., Moll J., Ponta H.,
 RA Hart I.R., Herrlich P.;
 RT "Problems with RHAMM: a new link between surface adhesion and
 RT oncogenesis?";
 RL Cell 95:591-592 (1998).
 CC -!- FUNCTION: INVOLVED IN CELL MOTILITY. WHEN HYALURONAN BINDS TO
 CC HMMR, THE PHOSPHORYLATION OF A NUMBER OF PROTEINS, INCLUDING THE
 CC FOCAL ADHESION KINASE OCCURS. MAY ALSO BE INVOLVED IN CELLULAR
 CC TRANSFORMATION AND METASTASIS FORMATION, AND IN REGULATING
 CC EXTRACELLULAR-REGULATED KINASE (ERK) ACTIVITY.
 CC -!- SUBUNIT: SUBUNIT OF THE HARC COMPLEX.
 CC -!- SUBCELLULAR LOCATION: CELL SURFACE AND CYTOPLASMIC.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=RHAMM1V4;
 CC IsoId=Q00547-1; Sequence=Displayed;
 CC Name=RHAMM1;
 CC IsoId=Q00547-2; Sequence=VSP_004287;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

FT CONFLICT 769 769 F -> L (IN REF. 1).
FT CONFLICT 787 787 Q -> E (IN REF. 1).
FT CONFLICT 881 881 Q -> E (IN REF. 1).
FT CONFLICT 907 909 HLL -> QJQ (IN REF. 1).
FT CONFLICT 920 920 G -> E (IN REF. 1).
FT CONFLICT 929 929 C -> Y (IN REF. 1).
SQ SEQUENCE 1690 AA; 189063 MW; D6F7916A9C532F16 CRC64;

Query Match 15.8%; Score 189; DB 1; Length 1690;
Best Local Similarity 27.1%; Pred. No. 0.014;
Matches 75; Conservative 54; Mismatches 106; Indels 42; Gaps 11;

QY 2 EKYDSVQSLVDVTAQFESYKALTASEIEDLKLENSLOEKAA---KAGKNEDVQHIL 58
Db E0IRELNQDLDEVITQNVKQADSSALDMLKQEGTEKSTLLEKTEKELVQSKQEA 688
QY 59 APESSNQEVYRMLDLOTKSALKETKEITVTSFLOKI-----TDLQNLKQOE-EDPR 111
Db 689 KTLNKEQLEKISDLK-QLAQEKLVREMTENAINQIQLEKESIEQQLALKQNELEDPO 747
QY 112 KOLEBERG---KAEKENTTAEITENKRWLLYBELYNKT---KPFQIQLDAPFEVKQ 164
Db 748 KKQSESEVHLOIKAQNTQDFELVESGESLKKLQQLQEQKTLGHEKLALEELKKEKE 807
QY 165 ALLNEHGAQEQOL-NKIRDSYAKL---LGHQNLKQKI-----KHVVKLKDNSOLK 211
Db 808 TIIEKEQELQOLQSKASAESALKVVQVLQOLQOQAAAASGEBSKTVAKUHDSEISQIK 867
QY 212 SEVSKLRLOL-----AKKQOSET---KLOEELNK 237
Db 868 SQAETSQELKSTQSNLEAKSKQLEAANGSLEEAKE 904

RESULT 6
USOL YEAST STANDARD; PRT; 1790 AA.
ID USOL1 YEAST
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intracellular protein transport protein USOL.
GN USOL1 OR INT1 OR YDL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Iwasaki M.;
RT "A cytoskeleton-related gene, usol1, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RN J. Cell Biol. 113:245-260(1991).
RN [2]
RP SEQUENCE OF 782-1790 FROM N.A.
RA Hoscutter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-8 FROM N.A.
RA Bai Y., Symington L.S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -!- SIMILARITY: BELONGS TO THE VDP/USOL1/YEL047C FAMILY.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X54378; CAA38253.1; --
EMBL; L03188; AAB00143.1; --
EMBL; U53668; AAB66659.1; --
SGD; S0002216; USOL1.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR006955; USOL_p115_C.
DR InterPro; IPR006953; USOL_p115_head.
DR Pfam; PF04871; USOL_p115_C; 1.
DR Pfam; PF04869; USOL_p115_head; 1.
KW Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
FT DOMAIN 1 724 GLOBULAR HEAD.
FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
FT CONFLICT 847 847 G -> E (IN REF. 2).
FT CONFLICT 924 924 E -> K (IN REF. 2).
FT CONFLICT 1253 1253 V -> I (IN REF. 2).
FT CONFLICT 1319 1319 I -> V (IN REF. 2).
FT CONFLICT 1461 1461 N -> S (IN REF. 2).
FT CONFLICT 1581 1581 G -> S (IN REF. 2).
FT CONFLICT 1600 1600 I -> V (IN REF. 2).
FT CONFLICT 1661 1661 R -> S (IN REF. 2).
FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).
SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E39FD4818 CRC64;

Query Match 15.5%; Score 185.5; DB 1; Length 1790;
Best Local Similarity 25.3%; Pred. No. 0.022;
Matches 72; Conservative 53; Mismatches 101; Indels 59; Gaps 10;

QY 5 DSMVQSLVDVTAQFESYKALTAS-----EIEDLKLENSLOEKAAKAGNAEDVQH 57
Db 1148 ESLEKEHEDLAAQLKYEEQIANKEQVNEIISQINDIITSTQENESIKKKNDLEGEV 1207
QY 58 LATESSNQEVYRM-----LLDLOTKSALKETKEITV-----SFLOKIDTLQNL 103
Db 1208 KAMKSTSEEQNLKKSEIDALNLQIKELKKNETNEASLESISKSVESVTVKIKELQDEC 1267
QY 104 KQOEEDPRKOLEDEGRKAEKENTTAELTEINKWRLLYBELYNKTPFQIQLDAF---- 159
Db 1268 NFKEKEV-SELEDKIKASEDKNSKYLELQKESK---IKEELDAKTTKIQLEKITVLS 1323
QY 160 -----EVE-----KQALNEHGAQEQNLKIR-----DSYAKLLGH-----QNL 193
Db 1324 KAKEKSELSRLKTSSEERKNAEQLEKLNKBIQKNQAFKERKLNESGSTITQEY 1383
QY 194 KQIKK-----HVVKLKDNSOLKSEVKRLCOLAKKQSEYKQOE 234
Db 1384 SEKINTLEDELIRLQNELKAKBIDNTRSELEKVSLSNDELLEE 1428

RESULT 7
MYH4 RABIT
ID MYH4 RABIT STANDARD; PRT; 1938 AA.
AC Q28641;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myosin heavy chain, skeletal muscle, juvenile.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]

```

RC STRAIN=New Zealand white; TISSUE=Skeletal muscle;
RA Maeda K., Hostinova E., Rosesch-Kleinlauf A., Schuster H., Gasperik J.,
RA Wittinghofer A.;
RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit
RT skeletal muscle and a novel cosynthesis of S-1 fragment with the
RT essential and regulatory light chains.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32574; AAA74199.1; -.
CC PIR; A59293; A59293.
CC HSP; P13538; 2MYS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC PRODOM; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MSC; 1.
CC PROSITE; PS50096; IQ; 1.
CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC Calmodulin-binding; ATP-binding; Methylation; Alkylation;
CC Multigene family.
CC DOMAIN 1 783 MYOSIN HEAD-LIKE.
CC FT DOMAIN 784 813
CC FT DOMAIN 842 1398
CC FT NP_BIND 179 186
CC FT DOMAIN 658 680
CC FT DOMAIN 760 774
CC FT MOD_RES 35 35 METHYLATION (MONO-) (BY SIMILARITY).
CC FT MOD_RES 120 120 METHYLATION (TRI-) (BY SIMILARITY).
CC FT MOD_RES 552 552 METHYLATION (TRI-) (BY SIMILARITY).
CC FT MOD_RES 756 756 METHYLATION (SH-1) (BY SIMILARITY).
CC FT MOD_RES 698 698 ALKYLATION (SH-2) (BY SIMILARITY).
CC FT MOD_RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).
CC SEQUENCE 1938 AA; 223064 MW; D8A8A2EC5B182626 CRC64;
Query Match 15.3%; Score 182.5; DB 1; Length 1938;
Best Local Similarity 22.5%; Pred. NO. 0.033;
Matches 81; Conservative 54; Mismatches 96; Indels 129; Gaps 13;
QY 8 VQSLDVTQAQESYKALTAIEDLKLNSLSQEKAAKGNAR-----DVQHQIILATESS 63
Dd 1210 IDNLRVQKLEKEKSELKMLDIDL-----ASNMTVSKAGNLEKMCRTLEDQVSELKTK 1265
QY 64 NOEYVRMLDLQTKSALKETBIKET-----VSFLQKITDLQNLQKQ 106

```

```

Db 1266 EEEHQLINDLSAQRARLQTESGFSSRLDEKDSLVLSQSGKQAFQTOIEELKQLBEE 1325
QY 107 -----BED---FRKQLEDEGRKAKENTTABLTETINKWRLLY----- 142
Dd 1326 IKAKSALAHALQSAHDCDLLEQVEEVEEQAKELQRAMSKANGSEVAQWRKYETDAQR 1385
QY 143 -EELYNTKTPFOIOLD-----AFVEVEKQALLNE-----HGNAQEQOL 177
Dd 1386 TEELEAKKLAQRLQDAEHEVAVNAKCALEKTKORLQNEVEDLMIDVERTNAACAL 1445
QY 178 NKIRDSVAKLIG-----HQLKQ-----KIKVVK-----LKDEN 207
Dd 1446 DKORNPDKILAEWKHKVEETHAELESQKSRSLSTEVFKVKNAYESLDQLETKEEN 1505
QY 208 SOLKSEVSKLRCQAK-----KKQ-----SETKQLELANKVLGIK 242
Dd 1506 KNLQOEISDLTEQIABGGKRIHELEKVKQVEQKSELQAALEBAEASLEHEGKILAIQ 1565
RESULT 8
MYH3 RAT
ID MYH3 RAT STANDARD; PRT; 1940 AA.
AC P12847;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=87060988; PubMed=3783701;
CC Strehler E.E., Strehler-Page M.-A., Perriard J.C., Periasamy M.,
CC Nadal-Ginard B.;
CC "Complete nucleotide and encoded amino acid sequence of a mammalian
CC myosin heavy chain gene. Evidence against intron-dependent evolution
CC of the rod.";
CC J. Mol. Biol. 190:291-317(1986).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04267; CAA27817.1; -.
CC PIR; A24922; A24922.
CC HSP; P13538; 2MYS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR004009; Myosin_N.

```


Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
Calmodulin-binding; ATP-binding; Methylation; Alkylation;
Multigene family.

CC NEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U87231; AAB47555.1; -;
CC EMBL; M16557; AAA48970.1; -;
CC PDB; 2MYS; 11-JAN-97.
CC PDB; 1ALM; 17-DEC-97.
CC PDB; 1M8Q; 10-SEP-02.
CC InterPro; IPR000048; IQ_region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR004009; Myosin_N.
CC InterPro; IPR002928; Myosin_tail.
CC Pfam; PF00612; IQ; 2.
CC Pfam; PF00663; myosin_head; 1.
CC Pfam; PF02736; Myosin_N; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PS00096; IQ; 1.
CC
CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;
CC Calmodulin-binding; Multigene family; 3D-structure.
CC
CC INIT MET 0
CC DOMAIN 1 782
CC DOMAIN 783 812
CC DOMAIN 812 840
CC DOMAIN 841 1938
CC NP_BIND 179 186
CC DOMAIN 657 679
CC DOMAIN 759 773
CC MOD_RES 1 1
CC MOD_RES 35 35
CC MOD_RES 130 130
CC MOD_RES 551 551
CC MOD_RES 755 755
CC MOD_RES 697 697
CC MOD_RES 707 707
CC CONFLICT 907 907
CC CONFLICT 980 980
CC CONFLICT 1343 1343
CC CONFLICT 1545 1545
CC CONFLICT 1796 1797
CC CONFLICT 1830 1830
CC CONFLICT 1863 1863
CC CONFLICT 1929 1931
CC TURN 7 8
CC TURN 11 12
CC TURN 14 17
CC TURN 21 24
CC TURN 25 28
CC TURN 34 36
CC TURN 37 41
CC STRAND 47 55
CC STRAND 59 62
CC STRAND 70 73
CC STRAND 74 76
CC STRAND 77 78
CC TURN 83 84
CC TURN 90 90
CC STRAND 91 93
CC
CC MYOSIN HEAD-LIKE.
CC IQ.
CC HINGE.
CC COILED COIL (POTENTIAL).
CC ATP (POTENTIAL).
CC ACTIN-BINDING.
CC ACTIN-BINDING.
CC ACETYLATION.
CC METHYLATION (MONO-).
CC METHYLATION (TRI-).
CC METHYLATION (TRI-).
CC METHYLATION (MONO-).
CC ALKYLATION (SH-1).
CC ALKYLATION (SH-2).
CC C -> Q (IN REF. 7 AND 8).
CC L -> F (IN REF. 1).
CC E -> D (IN REF. 5).
CC S -> A (IN REF. 5).
CC HV -> QL (IN REF. 5).
CC S -> A (IN REF. 5).
CC I -> V (IN REF. 10).
CC IHG -> FH (IN REF. 10).

FT HELIX 99 108
FT TURN 109 113
FT STRAND 116 118
FT STRAND 123 126
FT HELIX 133 135
FT TURN 137 138
FT HELIX 139 142
FT TURN 143 145
FT TURN 148 149
FT HELIX 155 169
FT TURN 170 170
FT STRAND 173 179
FT TURN 181 182
FT HELIX 185 199
FT TURN 200 200
FT HELIX 218 233
FT STRAND 234 235
FT STRAND 243 244
FT STRAND 247 254

Query Match 15.0%; Score 179.5; DB 1; Length 1938;
Best Local Similarity 20.5%; Pred. No. 0.047;
Matches 73; Conservative 55; Mismatches 107; Indels 121; Gaps 9;

QY 8 VQSLDYTAQPSYKALTASIEDLKLENSLOEYKAAKNAEDVQHCILATESNQBY 67
Db 1209 IDNLRVQKLEKESEKMEIDDLASNNESVSKANLEKNCRTLEDQLSIKTKEQN 1268
QY 68 VMMLDLQTSALKETEI-----KEITVS-----FLOKITDQNLKQOE--- 107
Db 1269 QRMINDINTQRLQTETGEYSRQAEKDALISQLSRGQGTQQLBELKRLHEEIKAK 1328
QY 108 -----BDFRKLQLEDEGRKAENKNTTAEITTEINKWRLY-----EEL 145
Db 1329 NALAHASARHDCLELLREQYEEQEAQKGLQALSKANSEVAQWRTKYETDAIQTEEL 1388
QY 146 YNKTQPFQIOLD-----APEVKKQALLNE-----HGAQOEQLNKR 181
Db 1389 BEAKKKLAQRLODAEEHVEVNAKASLEKTKQRLONEVDLMVDVERSNAAALDDKKQ 1448
QY 182 DSYAKLGHQNLK-----QKIKHVVKLKDNSQIK 211
Db 1449 KNFDKILAEWKQKYETQTELEASQKESRSLSTELFKMKNAVEESLDHLETKRENKQLQ 1508
QY 212 SEVSKLRQOLA-----KK-----KQSETKLOEELNKVLGIX 242
Db 1509 QEIADLTQIAEGGKAVHELEKVKKHVEQEKSELOASLEAEASLEHBEGLIRLQ 1564

RESULT 11
MYH4_HUMAN
ID MYH4_HUMAN STANDARD; PRT; 1939 AA.
AC Q9Y623;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, fetal (Myosin heavy chain IIB)
DE (MyHC-IIB).
GN MYH4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HEBICAL COILED COILS.
 -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMN). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 -!- SIMILARITY: Contains 1 myosin-like globular head domain.
 -!- SIMILARITY: Contains 1 IQ domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AF111783; AAD29949.1; -
 HSHP; P13538; 2MYS.
 Genew; HGNC:7574; MYH4.
 MIM; 160742; -
 InterPro; IPR000848; IQ region.
 InterPro; IPR001609; myosin head.
 InterPro; IPR004009; Myosin_N.
 InterPro; IPR002928; Myosin_tail.
 Pfam; PF00612; IQ; 2.
 Pfam; PF00063; myosin head; 1.
 Pfam; PF02736; Myosin_N; 1.
 Pfam; PF01576; Myosin_tail; 1.
 PRINTS; PR00193; MYOSINHEAV.
 ProDom; PD000355; myosin_head; 1.
 SMART; SM00015; IQ; 1.
 SMART; SM00242; MYSC; 1.
 PROSITE; PS0096; IQ; 1.
 Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 Multigene family.
 DOMAIN 1 784 MYOSIN HEAD-LIKE.
 DOMAIN 785 814 IQ.
 DOMAIN 843 1939 COILED COIL (POTENTIAL).
 NP_BIND 179 186 ATP (POTENTIAL).
 SEQUENCE 1939 AA; 223012 MW; 40BIAD1D777A47DE CRC64;

 Query Match 15.0%; Score 179.5; DB 1; Length 1939;
 Best Local Similarity 18.7%; Pred. No. 0.047; Indels 127; Gaps 7;
 Matches 67; Conservative 61; Mismatches 104;
 QY 8 VQSLEDVTAQFESYKALTAASHIEDLKLENSLSQEKAKAGNAEDVQHILATESSSNOY 67
 DB 1211 IDSQRVQKLEKEKSELKWEINDLASNETVSKAKANFERKMCRTLEDQLSEIKTEBEQ 1270
 QY 68 VRMLDLDTQKALKETEKEIT-----VSFLQKITDQLNQLKQOEED- 109
 DB 1271 QRLNELSAQKARLTESGEFSRQDEKDMVSQLSRGKQAFQTOIEBKRLQLEBETKAK 1330
 QY 110 -----FRKQLEDEGRKAEKENTTAELTEETINKWRLYE----- 143
 DB 1331 STLAHALQSARHDCDLLEQYEEBOEAKAELQRGKANSVAQFTQIETDAIQTIEL 1390
 QY 144 -----ELYNKTKPQIQIQL----- 156
 DB 1391 EBAKKKLAQRLODAEEHVEAVNSKASLEKTKQRLQNEVEDLMIDVRSNACIALDKKQ 1450
 QY 157 -----DAFEVEKQALLNEHGAQEQNLKIRDSYAKLLGH-----QNLK 194
 DB 1451 RNFDKVLAEWKQYETQAELEASQKESRSLSTELFKVKNAYEESLDHLETKRENKQLQ 1510

QY 195 OKI-----KHVVKLDENSQKSEYSKLRCLQAKKKQSETKLOEELNKVLGIK 242
 DB 1511 QEISDLTQIAEGGKHIELEKVKKQDHEKSELQTSL---EEAEASLEHEGKILRIQ 1566

 RESULT 12
 MFPL_LYCES
 ID MFPL_LYCES STANDARD; PRT; 697 AA.
 AC P93203;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE MAR binding filament-like protein 1.
 GN MFPL.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]_TaxID=4081;
 RP SEQUENCE FROM N.A.
 RC STRAINS=cv. VFNT Cherry; TISSUE=Fruit;
 RX MEDLINE=97112038; PubMed=8953774;
 RA Meier I., Phelan T., Gruissem W., Spiker S., Schneider D.;
 RT "MFPL, a novel plant filament-like protein with affinity for matrix
 attachment region DNA.";
 RL Plant Cell 8:2105-2115(1996).
 CC -!- FUNCTION: BINDS DNA. INTERACTS WITH CHROMATIN VIA MATRIX
 ATTACHMENT REGIONS (MARS). LIKELY TO PARTICIPATE IN NUCLEAR
 ARCHITECTURE BY CONNECTING CHROMATIN WITH THE NUCLEAR MATRIX AND
 POTENTIALLY WITH THE NUCLEAR ENVELOPE.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR MATRIX.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; Y07861; CAA69181.1; -
 PIR; T07111; T07111.
 KW Nuclear protein; DNA-binding; Coiled coil.
 FT DOMAIN 125 681 COILED COIL (POTENTIAL).
 FT DOMAIN 10 15 POLY-SER.
 SQ SEQUENCE 697 AA; 79516 MW; 700A56D68D6A7E49 CRC64;

 Query Match 15.0%; Score 179; DB 1; Length 697;
 Best Local Similarity 25.6%; Pred. No. 0.017; Indels 56; Gaps 12;
 Matches 72; Conservative 57; Mismatches 96;
 QY 1 QEKYDSMVQSLEDVTAQFESYKALTAS----EIEDLKLENSLSQEKAKAGNAEDVQH-- 55
 DB 427 QESLENSRSESDITVQLEQLDLSSKLERVSKLQME---LETRASLQRIDETKESS 483
 QY 56 QILATE-----SSNQEVVRMLDLDTQKALKETEKEITVSFLQKITDQLNQLKQ 106
 DB 484 ELLAAELTTTKELLKKTNEEMHTMSDELAVSENRSLSQTEL-VNVYKREHTRNELKQE 542
 QY 107 -----PEDFRKOLEDEEGKAEKENTTAELTEETINKWRL-LYEEL 145
 DB 543 KTVIRTLLEEELKFLSESITREKELRKSLDE---LEK---ATESLDEINRNVLALBEL 595
 QY 146 YNKTPTQIQIDAPEVEKEKQALLNEHGAQEQNLKIRDSYAKLL---CHQNLKQIKHV 201
 DB 596 ELATSRNSSLEDEVRHQSVSEBQKQISQAEQMLEDAHSLVMKLGKRESLEKRAK--- 652
 QY 202 KLDENSQKSEYSKLRCLQAKKKQSETKLOEELNKVLGIK 242
 DB 653 KLEDEMAAAKGEILLRLRSQINSVK---APVEDEEKVVAGEK 690

RESULT 13

MYH1 HUMAN
 ID MYH1 HUMAN STANDARD; PRT: 1939 AA.
 AC P12882; Q9Y622;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, skeletal muscle, adult 1 (Myosin heavy chain
 DE IIX/d) (MYHC-IIX/d).
 GN MYH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=99318869; PubMed=10388558;
 RA Weiss A., Schiaffino S., Leinwand L.A.;
 RT "Comparative sequence analysis of the complete human sarcomeric myosin
 RT heavy chain family: implications for functional diversity.";
 RL J. Mol. Biol. 290:61-75 (1999).
 RN [2]
 RP SEQUENCE OF 1064-1939 FROM N.A.
 RX MEDLINE=86176778; PubMed=2421254;
 RA Saez L., Leinwand L.A.;
 RT "Characterization of diverse forms of myosin heavy chain expressed in
 RT adult human skeletal muscle.";
 RL Nucleic Acids Res. 14:2951-2969 (1986).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF111785; AAD29951.1; -
 DR EMBL; X03740; CAA27380.1; -
 DR F1; A23767; A23767.
 DR HSP; P13538; 2MYS.
 DR Genew; HGNC:7567; MYH1.
 DR MIM; 160730; -
 DR InterPro; IPR000048; IQ region.
 DR InterPro; IPR001609; myosin head.
 DR InterPro; IPR004009; myosin N.
 DR InterPro; IPR002928; Myosin tail.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF00063; myosin head; 1.
 DR Pfam; PF02736; Myosin N; 1.
 DR Pfam; PF01576; Myosin tail; 1.
 DR PRINTS; PR00183; MYOSINHEAVY.
 DR PRODom; PD000355; myosin head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS50096; IQ; 1.

Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 Multigene family.
 KW MYOSIN HEAD-LIKE.
 KW DOMAIN 1 784
 KW DOMAIN 785 814
 KW DOMAIN 843 1939
 KW NP_BIND 179 186
 KW DOMAIN 659 681
 KW DOMAIN 761 775
 KW MOD_RES 130 130
 KW MOD_RES 699 699
 KW MOD_RES 709 709
 KW CONFLICT 1131 1131
 KW CONFLICT 1139 1139
 KW CONFLICT 1158 1158
 KW CONFLICT 1163 1163
 KW CONFLICT 1286 1289
 KW CONFLICT 1302 1303
 KW CONFLICT 1451 1451
 KW CONFLICT 1470 1470
 KW CONFLICT 1473 1474
 KW CONFLICT 1569 1569
 KW CONFLICT 1598 1598
 KW CONFLICT 1606 1606
 KW CONFLICT 1643 1643
 KW CONFLICT 1648 1648
 KW CONFLICT 1750 1750
 KW CONFLICT 1822 1822
 KW CONFLICT 1845 1845
 KW SEQUENCE 1939 AA; 223114 MW; 39ADB26AB79DFA53 CRC64;
 Query Match 14.9%; Score 178.5; DB 1; Length 1939;
 Best Local Similarity 23.3%; Pred. No. 0.052;
 Matches 84; Conservative 51; Mismatches 96; Indels 129; Gaps 14;
 QY 8 VOSLEDVTAQFSYKALTAASELDKLENSFQEKAAKAGKNAED---VQHQILATES 63
 DB 1211 INLGRVQKLEKSEKMKQEIDDL-----ASNETVSKAKGNLKNCRALDQLSEIKTK 1266
 QY 64 NOEYVRLDLDTQKALKETET-----KEITVS-----FLOKITDLQNLKQK 106
 DB 1267 BEEQQLINDLTAQARLCTESGEYSRQDLEKDLVLSQSRGKQAFQIQIEELKRLLEE 1326
 QY 107 -----BED---FKQLEDEBGRKAEXNTTAEITTEINKWRLLY----- 142
 DB 1327 IKAKSALAHALOSSRHDCDLLREQYEEBQKAEIQRAMSKANSEVAQWRTKYETDAIGR 1386
 QY 143 -EELYNKTKPFQIQLD-----APEVEKQALLNE-----HGAAQEQ 177
 DB 1387 TEELEBAKKKLAQRLODAEEHVEAVNAKASLEKTKQRLQNEVEDLMIDVERTNAACAL 1446
 QY 178 NKIRDSYAKLLG-----HQLNKQ-----KIKHVVK-----LKDSN 207
 DB 1447 DKQRNFKILAEWKQKCEETHAELEASQKESRSLSLTFKIKWAYEBSLDQLETKREN 1506
 QY 208 SOLKSEVSKRLQRLAK-----KKQ-----SETKQELNKLVLGK 242
 DB 1507 KNLQEIISDLTEQIAEGGKRIHELEKIKKQVEKSELQAALEAEASLEHEEGKILRIQ 1566
 RESULT 14
 MYH3 HUMAN
 ID MYH3 HUMAN STANDARD; PRT: 1940 AA.
 AC P11055; Q15492;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic
 DE myosin heavy chain) (SMHCE).
 GN MYH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263803; PubMed=2726495;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H.,
RA Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Nucleotide sequence of full length human embryonic myosin heavy
chain cDNA.";
RL Nucleic Acids Res. 17:3591-3592(1989).
RN [2]
RP SEQUENCE OF 774-1940 FROM N.A.
RX MEDLINE=90033298; PubMed=2806546;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,
RA Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Human embryonic myosin heavy chain cDNA. Interspecies sequence
conservation of the myosin rod, chromosomal locus and isoform
specific transcription of the gene.";
RL FEBS Lett. 256:21-28(1989).
RN [3]
RP SEQUENCE OF 856-1940 FROM N.A.
RX TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bobber E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [4]
RP SEQUENCE OF 856-1940 FROM N.A.
RX MEDLINE=89366648; PubMed=2771643;
RA Karsch-Mizrachi I., Travis M., Blau H., Leinwand L.A.;
RT "Expression and DNA sequence analysis of a human embryonic skeletal
muscle myosin heavy chain gene.";
RL Nucleic Acids Res. 17:6167-6179(1989).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE
AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL
MUSCLE.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X13988; CAA32167.1; -
CC EMBL; X13100; CAA31492.1; -
CC EMBL; X51593; CAA35942.1; -
CC EMBL; X15696; CAA33731.1; -
CC PIR; S04090; S04090.
CC HSSP; P13538; 2MYS.
CC Genew; HGNC:7573; MYH3.
CC MIM; 160720; -
CC GO; GO:0007517; P:muscle development; TAS.
CC InterPro; IPR000048; IQ region.
CC InterPro; IPR001609; myosin_head.
DR

```

```

DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 556 578 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
FT CONFLICT 1331 1331 A -> G (IN REF. 3).
FT CONFLICT 1391 1392 KK -> QE (IN REF. 1 AND 2).
FT CONFLICT 1608 1609 SR -> RA (IN REF. 3).
FT CONFLICT 1663 1664 RG -> QT (IN REF. 2).
SQ SEQUENCE 1940 AA; 224035 MW; 43CA58C6A4BA1253 CRC64;

Query Match 14.9%; Score 178.5; DB 1; Length 1940;
Best Local Similarity 19.1%; Pred. No. 0.052;
Matches 68; Conservative 62; Mismatches 105; Indels 121; Gaps 7;

QY 8 VQSLIEDVTAQPSYKALTASIEDLKLENSLSOEKAKAGNAEDVQHOILATESNOEY 67
Db 1208 IDNLRVQVKLEKSEFKLEIDLLSSMESVSKANLEKICRTLEDQSEARQKNEI 1267
QY 68 VRMLLDQTSKALKEITEIT-----VSFLOKITDLQNLQKQDEED- 109
Db 1268 QRSLSLTQSRLOTEBAGELSRLQSEKESIVSLSRSKQFTQOTBELKQLEENKAK 1327
QY 110 -----PRKQLEDEEGKAEKENTAEITTEINKWRLIYE----- 143
Db 1328 NALAHQLSSRHDCDLLREQVEEYEEQEGKAEQLQALSKANSEVAQWRTKYETDAIQTEEL 1387
QY 144 -----ELYN-----KTK----- 150
Db 1388 EAACKLAQRLODSEEQEAVNAKCAKSLKTKQRLQGEVEDLMTDVERANSLAALDKKQ 1447
QY 151 -----PFQIQIDAFVEKQALLNEHGAQEQINKIRDSYAKLLG-----HQLK 194
Db 1448 RNFQKVLAEWTKCEESQAELEASLSEKSRSLSTELFKLNAYEALDQLETVREKNKLE 1507
QY 195 QKIKHVKLKXDNESQKASEVSKLRCOLAKK-----QSEYKLOELKVLGIK 242
Db 1508 QETADLTQEAENGKTIHELEKSKKQIELEKADTQIALALEAEAEALAEHEEAKILFIQ 1563

RESULT 15
SCPI_MESAU
ID SCPI_MESAU STANDARD; PRT; 845 AA.
AC Q60563;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Synaptonemal complex protein 1 (SCP-1 protein) (Meiotic chromosome
synaptic protein) (fragment).
GN SCPI OR SYNI.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;

```



```

CC
DR EMBL; M36769; AAC17185.1; --
DR EMBL; Z38133; CAA86239.1; --
DR EMBL; X51592; CAA35941.1; --
DR EMBL; AF067143; AAC211557.1; --
DR PIR; I38055; I38055.
DR HSSP; P13538; 2MYS.
DR Genew; HGNC:7578; MYH8.
DR MIM; 160741; --
DR GO; GO:0005859; Cmuscle myosin; TAS.
DR GO; GO:0008307; F:structural constituent of muscle; TAS.
DR InterPro; IPR000408; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00663; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00996; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding. 780
FT DOMAIN 1 MYOSIN HEAD-LIKE.
FT DOMAIN 781 813
FT DOMAIN 842 1937
FT NP_BIND 181 188
FT DOMAIN 658 680
FT DOMAIN 760 774
FT MOD_RES 132 132
FT MOD_RES 698 698
FT MOD_RES 708 708
FT CONFLICT 15 15
FT CONFLICT 970 970
FT CONFLICT 1072 1072
FT CONFLICT 1247 1247
FT CONFLICT 1251 1252
FT CONFLICT 1261 1261
FT CONFLICT 1297 1297
FT CONFLICT 1377 1378
FT CONFLICT 1504 1505
FT CONFLICT 1847 1847
FT CONFLICT 1914 1914
SQ SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E98 CRC64;
Query Match 14.8%; Score 176.5; DB 1; Length 1937;
Best Local Similarity 22.8%; Pred. No. 0.066;
Matches 84; Conservative 53; Mismatches 97; Indels 135; Gaps 13;
QY 5 DSM-----VQSLDVTAFQESYKALTASEIEDLKLNSLSLOEKAAGKNAE-----DVO 54
DB 1201 DSMALGQIDNLQVQKLEKEKELKMETDLDL-----SSNAEATSKAGNLEKMCRSLE 1256
QY 55 HOILATESNQBYRVLMLDLOPKSALKETEI-----KEITVSFL-----QKIT 97
DB 1257 DOVSELKTKKEEQRLINDLTACRARLQTEAGESYRQLDEKDALVSQLSRKQASTQOIE 1316
QY 98 DLONOLKQOED-----FRQLEDEGRKAEKENTTAELTEERINKWEL 140
DB 1317 ELKHQLEETKAKNALAHALQSSRHDCDLLRLRQYEEBQEGKAEORALSKANSEVAQWRT 1376
QY 141 LY-----EELNNTKFPQIQLD-----AFVEKQALINE----- 169
DB 1377 KYETDAIQTEELEAKKLAQLQAEHVEAVNAKASLEKTKQRLQNEVEDMLDVE 1436
QY 170 -HGAQEQINKTRDSYAKLLGHONLKQ-----KIKHVVK----- 202
DB 1437 RSNAACAALDKQRNFDKVLSEWQKYEETQAELEASQKESRSLSTELFKVKNVYEESLD 1496

```

```

QY 203 ----LKDENSQKSVSKLRQLAQ-----KKQ-----SETKQAE 233
DB 1497 QLETLRENNKLNQQRISDLTEQIARGGKQIHELEKIKQVEQKCEIQAALEAEASLEH 1556
QY 234 ELNKLVLGK 242
DB 1557 BEGKILRIQ 1565

RESULT 17
MYH3 CHICK
ID MYH3 CHICK STANDARD; PRT; 1940 AA.
AC P02555;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87194881; PubMed=3571266;
RA Molina M.I., Kropp K.E., Gulick J., Robbins J.;
RT "The sequence of an embryonic myosin heavy chain gene and isolation
RT of its corresponding cDNA.";
RL J. Biol. Chem. 262:6478-6488(1987).
RN [2]
RP SEQUENCE OF 1502-1940 FROM N.A.
RX MEDLINE=8161144; PubMed=6833296;
RA Kavinsky C.J., Umeda P.K., Sinha A.M., Elzinga M., Tong S.W., Zak R.,
RA Jakovcic S., Rabinowitz M.;
RT "Cloned mRNA sequences for two types of embryonic myosin heavy chains
RT from chick skeletal muscle. I. DNA and derived amino acid sequence of
RT light meromyosin.";
RL J. Biol. Chem. 258:5196-5205(1983).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DB EMBL; V00430; CAA23712.1; --
DR EMBL; J02714; AAA48972.1; --
DR PIR; A29320; A29320.
DR HSSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.

```

DR Pfam; PF00063; myosin head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_Tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS00096; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
 FT DOMAIN 786 815 IO.
 FT DOMAIN 844 1940 COILED COIL (POTENTIAL).
 FT NP BIND 179 186 ATP.
 FT DOMAIN 560 682 ACTIN-BINDING.
 FT DOMAIN 762 776 ACTIN-BINDING.
 FT MOD RES 130 130 METHYLATION (TRI-) (POTENTIAL).
 FT MOD RES 700 700 ALKYLATION (SH-1).
 FT MOD RES 710 710 ALKYLATION (SH-2).
 FT VARIANT 379 379 G -> D.
 FT CONFLICT 1547 1547 T -> A (IN REF. 2).
 FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
 FT SEQUENCE 1940 AA; 222816 MW; C34833D75B04DF2 CRC64;
 Query Match 14.8%; Score 176.5; DB 1; Length 1940;
 Best Local Similarity 20.2%; Pred. No. 0.066;
 Matches 72; Conservative 53; Mismatches 110; Indels 121; Gaps 9;
 QY 8 VQSLDVTAQFESYKATASEIEDLKLENSLOEKAAGKAGNAEDVQHILATESNOEY 67
 Db 1212 IDNLQVQKLEKSEKSLKMEIDDLASNMESVSKANLEKMCESLQSLSEIKTEEEQ 1271
 QY 68 VRMLDLQTSALKETHEIKIT-----VSFQKITDQNLQKQ----- 106
 Db 1272 QRTINDISAQKARLQTESGYSRQVEEKDALISQLSRGKQAFQOIBELKRLHEEITKAK 1331
 QY 107 -----EED-----FRKQLEBEGKKAENTTAELTEINKWLLY-----EEL 145
 Db 1332 KCPAHALQSARHCDLLREQVEEQAQKGLQALSKANSEVAQWRKYETDAIQTEEL 1391
 QY 146 YNKTQFQIOLD-----APEVEKQALLNE-----HGAQEQQLNKIR 181
 Db 1392 EEAQKLAQRLQDAEEHVEAVNGKASLEKTKQLQNEVEDLMIDVRSNAACALDKKQ 1451
 QY 182 DSVALKLGHNK-----OKIKHVVKLKDENSOLK 211
 Db 1452 KNFDKILSEWKQYEETQAEASQXESRSLSTELPFMKNAYEESLDHLETLKREKNLQ 1511
 QY 212 SEVSKLRCQLAK-----KKQ-----SETKLOEELNKVLGK 242
 Db 1512 QEISDLTEQIAEGGKAIHELEKVKQIEQKSELQTALEAEASLSHEGKILRVQ 1567
 RESULT 18
 SCPI MOUSE
 ID SCPI MOUSE STANDARD; PRT; 993 AA.
 AC Q62209; Q09205; P70192; Q62329;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Synaptonemal complex protein 1 (SCP-1 protein).
 GN SYCP1 OR SCPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC STRAIN=CBA; TISSUE=Testis;
 RX MEDLINE=96004899; PubMed=7548215;
 RA Sage J., Martin L., Cuzin F., Rassoulzadegan M.;
 RT "cDNA sequence of the murine synaptonemal complex protein 1 (SCP1).";

RL Biochim. Biophys. Acta 1263:258-260 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss; TISSUE=Testis;
 RA Kerr S.M., Taggart M.H., Lee M., Cooke H.J.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-149 FROM N.A.
 RC STRAIN=C57BL/6;
 RA Sage J., Li Y., Martin L., Mattei M.-G., Guenet J.-L., Liu J.G.,
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 95-787 FROM N.A.
 RC STRAIN=ICR; TISSUE=Testis;
 RA Tsuchida J., Nishina Y., Nozaki M., Uchida K., Nishimune Y.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF
 CC SYNAPTONEMAL COMPLEXES (SCS), FORMED BETWEEN HOMOLOGOUS
 CC CHROMOSOMES DURING MEIOTIC PROPHASE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF
 CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.
 CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS
 CC IS FOUND TOWARDS THE CENTRE OF THE SYNAPTONEMAL COMPLEX WHILE THE
 CC C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE
 CC SYNAPTONEMAL COMPLEX (BY SIMILARITY).
 CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,
 CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL
 CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z38118; CAA86262.1; -
 DR EMBL; L41069; AAA64514.1; ALT_INIT.
 DR EMBL; U62864; AAC53335.1; -
 DR EMBL; U62860; AAC53335.1; JOINED.
 DR EMBL; U62861; AAC53335.1; JOINED.
 DR EMBL; U62862; AAC53335.1; JOINED.
 DR EMBL; U62863; AAC53335.1; JOINED.
 DR EMBL; D88539; BAA13639.1; -
 DR PIR; S49461; S49461.
 DR MGI; MGI:105931; Sycp1.
 KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
 KW DNA-binding; Coiled coil.
 FT DOMAIN 12 97
 FT DOMAIN 104 815 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 114 117 COILED COIL (POTENTIAL).
 FT DOMAIN 697 700 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 898 901 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 978 986 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT CONFLICT 527 527 ARG/LYS-RICH (BASIC).
 FT SEQUENCE 993 AA; 115962 MW; 1A4FA79D64FAF66 CRC64;
 Query Match 14.6%; Score 175; DB 1; Length 993;
 Best Local Similarity 21.3%; Pred. No. 0.04;
 Matches 73; Conservative 68; Mismatches 85; Indels 116; Gaps 14;
 QY 11 LEDVTAQFESYKALTASIEDLKLENSLOEKAAGKAGNAE----- 51
 Db 416 LEEMT-KFNNKEVELEKILLAEQKLDEKQVEKLAELQEKQELATFLLETRKE 474
 QY 52 --DVHQHILATESNQYVRMLDLQ-----KSKAKETIKETIWS-----FLOKI 96
 Db 475 VHDLQEQVTVTKTSEQHYLKQVEEMKTELEKLNKT---ELTASCDMLLENKFFVEA 531
 QY 97 TDLQNLKQOED-----FRKOLE-----DEEGRA 122

Db 532 SDVALEKKGQEDIIINCKQEEILLKQIENLEBKEMHLRDELSVRKEFIQQQDEVCKKL 591
 QY 123 EKENTTA-----ELTEINKWRL-----YELINKTKPFO-----153
 Db 592 DKSEENARSIECEVLKKEKQKILESCKNNLKQVENKSNIBELHQENKTLKKSSAEI 651
 QY 154 IQDAFEVEKQALINEHGAQEQOLNKIRDSYAKLGHONLKO--KIKHVVKLK--DENS 208
 Db 652 KQLNAYEIKVSKLELELESKQFEETNNYQEIENKLSBKGLLEVEKAKATVDEAV 711
 QY 209 OLKSEVSKLRQC-----LAKKQSETKLOFELNKVLGI 241
 Db 712 KLOKEID-LRQKHIAEMVMEKHKHOYDKIVEERDSELG 752

RESULT 19
 RASO_SULSO STANDARD; PRT; 864 AA.
 AC Q97WH0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR SSO2249.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic V., Allard G.,
 RA Awayez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Koza C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Gartlett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
 CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE006829; AAK42417.1; -
 CC FIR; B90395; B90395.
 CC HAMAP; MF_00449; -; 1.
 CC InterPro; IPR003439; ABC transporter.
 CC InterPro; IPR003405; SMC C.
 CC InterPro; IPR003395; SMC N.
 CC Pfam; PF04423; Rad50_zn_hook; 1.
 CC Pfam; PF02483; SMC C; 1.
 CC Pfam; PF02463; SMC N; 1.
 CC ProDom; PD000006; ABC transporter; 1.
 KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
 FT NP BIND 30 37 ATP (BY SIMILARITY).
 FT DOMAIN 164 701 COILED COIL (POTENTIAL).
 FT SEQUENCE 864 AA; 101601 MW; 657076AEA9B709FC CRC64;

Query Match 14.4%; Score 172.5; DB 1; Length 864;
 Best Local Similarity 25.3%; Pred. No. 0.046;
 Matches 68; Conservative 51; Mismatches 99; Indels 51; Gaps 10;
 QY 2 EKYSVMVQSLSDVTAQFESYKALTASEIB-----DLKLENSLQ-EYAAKAGKNAE-----51
 Db 272 EKEINLEENLRNIKLFKEIKYEVLAKSHTSMNSANVINLEKEIEYEXAIRRKESELPYKLK 331
 QY 52 --DVHQIILATSSNQYVVRMLDLDTQKSAKETEITKEITVSFLOKITDLOLQKQOEED 109
 Db 332 YAELEKRLBELQPKYQYLKSLKSLDLSKLNLERLEK--ASELSNDIDKVNLSLEKQVEE 389
 QY 110 FRKQLEDEGRKAQENTTAETAEITEINK-----WLLYEE-----LYNKTTPFP 153
 Db 390 TRKQQLNLFQAQAKVESLISEKNEIINNIQSVEGETCPVCGRPDLDEHKQKIIEAKSYI 449
 QY 154 IQDAFEVEKQALINEHGAQEQOLNKIRDSYAKL-----LGHONLKOQIKHVKVLDKDENSQ 209
 Db 450 LQJ-----ELNKNLELEELKKITNELANKIEREYRLSNKKASYDNVMKOLK---KLNSEIEN 503
 QY 210 LKSEVSKLRQCLAKKKQSETKLOBELNKV 238
 Db 504 LHSETESLK-----NIDEIKKI 521

RESULT 20
 MYSP_SCHMA STANDARD; PRT; 866 AA.
 AC P06198;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Paramyosin.
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
 OC Schistosomatidae; Schistosomatidae; Schistosoma.
 OX NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91270282; PubMed=2052029;
 RA Laclette J.P., Landa A., Arcos L., Willms K., Davis A.E.,
 RA Shoemaker C.B.;
 RT "Paramyosin is the Schistosoma mansoni (Trematoda) homologue of
 RT antigen B from Taenia solium (Cestoda).";
 RL Mol. Biochem. Parasitol. 44:287-296 (1991).
 RN [2]
 RP SEQUENCE OF 303-742 FROM N.A.
 RX MEDLINE=87018840; PubMed=3094144;
 RA Lanar D.E., Pearce E.J., James S.H., Sher A.;
 RT "Identification of paramyosin as schistosome antigen recognized by
 RT intradermally vaccinated mice.";
 RL Science 234:593-596 (1986).
 CC -!- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
 CC MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M35499; AAA29915.1; -
 CC EMBL; M4163; AAA29914.1; -
 CC InterPro; IPR002928; Myosin tail.
 CC Pfam; PF01576; Myosin tail; 1.
 CC Coiled coil; Muscle protein; Thick filament; Myosin.
 FT DOMAIN 1 22 NONHELICAL REGION (POTENTIAL).
 FT SEQUENCE 866 AA; 839 COILED COIL (POTENTIAL).

```

FT DOMAIN      840 866 NONHELICAL REGION (POTENTIAL).
FT DISULFID    750 750 INTERCHAIN (POTENTIAL).
FT CONFLICT    423 428 DOVEL > GSSORI (IN REF. 2).
FT CONFLICT    430 431 SS -> KL (IN REF. 2).
FT CONFLICT    637 637 V -> L (IN REF. 2).
FT CONFLICT    639 639 T -> I (IN REF. 2).
FT CONFLICT    691 691 G -> E (IN REF. 2).
FT CONFLICT    720 720 S -> F (IN REF. 2).
SQ SEQUENCE    866 AA; 100387 MW; 42FASBSE78176AE0 CRC64;

Query Match      14.4%; Score 172.5; DB 1; Length 866;
Best Local Similarity 22.6%; Pred. No. 0.046;
Matches 77; Conservative 50; Mismatches 96; Indels 117; Gaps 12;

QY 1 QEKVDSMVQSLQEDVTAQFESYKALT-----ASHIEDKLKNSLSQEKAAK 45
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 KRKFTWITELED-TAERERKKAVALSKLTKLTLEIKDLQSEISLENSLIRAKA 364
QY 46 AGKNAEDVQHQI-----LATSSQEVYVRLDLDQTSKALKETE-----84
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 ABSLASDQRRVEDLTIEVNTLSQNSQLESENRLKSLVNDLTDKNLLERENQNDQ 424
QY 85 IKETVSLF---OKITDQ-----NOL 103
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 VKELKSLRDANRLTDLEALRSQLEAERONLASALHDAEALHMDQKYQASQAALNHL 484
QY 104 K-QEEDFRKQLEDEGRKAENKTTAETLTKINKVRLLYE-ELYNKTRFQIQDAFEV 161
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 485 KSEWQRLRERDELSRLKSTTRTTEELFTVTITEMEVKYSLSRLKKRYESNIADLEI 544
QY 162 EKOALLNHEGAQEQLNKIRDSYAKLL-GHONLKQKIKHVVKLKDENS-----208
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 545 -----QLDTANKANLNMKENKNSQVKOLETFLDERRLRERAEANNLQI 590
QY 209 -----OLKSEVSKLRCQIAK-----KQSETKLOEBLNKY 238
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 591 TEKRLQLANEIEIRSTLENRLRKAETLEEEAQRV 630

RESULT 21
MYSS CYPCA
ID MYSS CYPCA STANDARD; PRT; 1935 AA.
AC Q90339;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]_TaxID=7962;
RP SEQUENCE FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=97352533; PubMed=9208928;
RA Hirayama Y., Watabe S.;
RT "Structural differences in the crossbridge head of temperature-associated myosin subfragment-1 isoforms from carp fast skeletal muscle.";
RT associated myosin subfragment-1 isoforms from carp fast skeletal muscle.";
RL Eur. J. Biochem. 246:380-387(1997).
RN [2]
RP SEQUENCE OF 981-1935 FROM N.A.
RC TISSUE=Fast muscle;
RX MEDLINE=97176447; PubMed=9023993;
RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;
RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal muscle and their gene expression associated with temperature acclimation.";
RL J. Exp. Biol. 200:27-34(1997).
RN [3]
RP SEQUENCE OF 1387-1528 FROM N.A.
RX MEDLINE=95194396; PubMed=7887920;

```

```

RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H.,
RA Uozumi T., Hirono I., Aoki T.;
RT "Temperature acclimation induces light meromyosin isoforms with
RT different primary structures in carp fast skeletal muscle.";
RL Biochem. Biophys. Res. Commun. 208:118-125(1995).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -1- SIMILARITY: Contains 1 IQ domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D89992; BAA22069.1; -
DR EMBL; D50476; BAA09069.1; -
DR EMBL; D43700; BAA07802.1; -
DR PIR; I50496; I50496.
DR HSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 812 839 HINGE.
FT DOMAIN 840 1935 COILED COIL (POTENTIAL).
FT NP_BIND 178 185 ATP (POTENTIAL).
FT DOMAIN 659 681 ACTIN-BINDING.
FT DOMAIN 761 775 ACTIN-BINDING.
FT MOD_RES 129 129 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
SQ SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;

Query Match      14.4%; Score 171.5; DB 1; Length 1935;
Best Local Similarity 21.7%; Pred. No. 0.12;
Matches 81; Conservative 55; Mismatches 102; Indels 135; Gaps 12;

QY 1 QEKVDSM-----VQSLQEDVTAQFESYKALTASIEDKLKNSLSQEKAAKAGNAE--- 51
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1195 KEQADSVAEELGEQIDNLQRVKQKLEKSEYKMEIDDL-----TSNMEAVAKAKALEQVC 1250
QY 52 -DVQHQILATSSNQEVYVRLDLDQTSKALKET-----EIKETVTS-----FL 93
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



```

DR PRINTS; PRO0193; MYOSINHEAVY.
DR ProDom; PD003376; DIL; 1.
DR ProDom; PD003355; myosin_head; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS00096; IQ; 3.
DR MYOSIN; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
KW Coiled coil.
FT DOMAIN 1 809 MYOSIN HEAD-LIKE.
FT DOMAIN 824 851 IQ 1.
FT DOMAIN 872 901 IQ 2.
FT DOMAIN 943 972 IQ 3.
FT DOMAIN 973 1812 COILED COIL (POTENTIAL).
FT DOMAIN 2060 2220 DILUTE.
FT NP BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 669 749 ACTIN-BINDING.
FT CONFLICT 191 191 L -> F (IN REF. 2).
FT CONFLICT 284 284 A -> T (IN REF. 2).
FT CONFLICT 291 291 G -> R (IN REF. 2).
FT CONFLICT 332 347 NKGCFEIEGVSDDEH -> IEWPFELKYRMKS (IN REF. 2)
FT CONFLICT 550 550 N -> K (IN REF. 2).
FT CONFLICT 865 866 HH -> QQ (IN REF. 2).
FT SEQUENCE 2245 AA; 258478 MW; 615E5EFD1AB45BE CRC64;

Query Match 14.0%; Score 167; DB 1; Length 2245;
Best Local Similarity 22.1%; Pred. No. 0.23;
Matches 56; Conservative 60; Mismatches 113; Indels 24; Gaps 6;

QY 2 EKYSWVSLDVTQAQFYSKALTASEIEDLKLENSLQKAAKAGNAEDVQHQLATE 61
Db 1118 QQLDPNKFQEFDRDNTNNQLEIQOLKANKSTLEEDYFSLGIRDNLERQVLELR 1177

QY 62 SSNQFVVMILD-----LQTKALKETEKEITVSLFKITDLQNLKQOEDFRQ 113
Db 1178 DENQ-LIKERLDSLQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQSQ 1236

QY 114 -----LEDEGRKAENKNTAEITTEINKWLLYELLYNKTFFQIOLDFAFEVKAL 166
Db 1237 INQLELTDHKSQIQQLQQLTEQNEKIKUKGLKEEYQDBKQLOQOELEIRKOSQSV 1296

QY 167 LNEHGAQEQNLKIRDSYAKL---LGHQNLKQIKKHVVVKLDENSQKSEVSKLRQOLAK 223
Db 1297 EDEKNSLTQLTTFVFESTQVSTNSHQ--KEKI---TTLKSTIELANKSIGKLAQEQKN 1351

QY 224 KKQSETKLOEIN 236
Db 1352 KQDEIRKIQFELN 1364

RESULT 25
MYSS_RABIT STANDARD; PRT; 1084 AA.
ID P02562;
AC P02562;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myosin heavy chain, skeletal muscle (Fragments).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
[1]
RP SEQUENCE OF 1-258.
RA Capony J.-P., Elzinga M.;
RT "The amino acid sequence of A 34,000 dalton fragment from S-2 of
RT myosin."
RL Biophys. J. 33:148A-148A(1981).
[2]
RP SEQUENCE OF 259-428.
RX MEDLINE=85131142; PubMed=3972832;
RA Lu R.C.; Wong A.;
RT "The amino acid sequence and stability predictions of the hinge

```

```

RT region in myosin subfragment 2.";
RL J. Biol. Chem. 260:3456-3461(1985).
[3]
RN SEQUENCE OF 409-1084 FROM N.A.
RX MEDLINE=87304245; PubMed=3305014;
RA Maeda K., Szczakiel G., Wittlinghofer A.;
RT "Characterization of cDNA coding for the complete light meromyosin
RT portion of a rabbit fast skeletal muscle myosin heavy chain.";
RL Eur. J. Biochem. 167:97-102(1987).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HNM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
CC CONSERVED.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X05958; CRA29391.1; -.
DR PIR; A02985; A02985.
DR PIR; A05280; A05280.
DR PIR; S00084; S00084.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Multigene family.
FT NON TER 1 1
FT DOMAIN <1 >258 ALPHA-HELICAL TAILPIECE (S2).
FT NON CONS 258 259
FT DOMAIN <259 1084 RODLIKE TAIL (S2 AND LMM DOMAINS).
FT DOMAIN 455 1084 COILED COIL (POTENTIAL).
FT VARIANT 405 405 L -> V.
FT VARIANT 408 408 V -> L.
FT VARIANT 421 421 E -> D.
FT VARIANT 423 423 S -> G.
FT VARIANT 426 426 K -> R.
FT SEQUENCE 1084 AA; 125488 MW; 229CFD69AGE1F7F0 CRC64;

Query Match 13.9%; Score 166.5; DB 1; Length 1084;
Best Local Similarity 18.9%; Pred. No. 0.12;
Matches 68; Conservative 61; Mismatches 103; Indels 127; Gaps 9;

QY 8 VQSLDVTQAFESYKALTASEIEDLKLENSLQKAAKAGNAEDVQHQLATSSNQY 67
Db 356 IDNLQRYVKQLEKEKSELKMEIDDLAGNMETVSKAGNLEKWCRTLEDQLSEVTKKEEH 415

QY 68 VMMLDLQTSKALKETIKIT-----VSFLQKITDLQNLKQOED- 109
Db 416 QRLNELSAQKARLHTSGEGFSLDEKDAWVQSLSGGQAFQIQLKQLEETKAK 475

QY 110 -----FRKQLEDEGRKAENKNTAEITTEINKWLL-----LYEEL 145
Db 476 SALAHALQSSRRDCDLLEQVEEQAELQKRAMKANSEVSWRYKCTDAIQTEEL 535

QY 146 YNKTQFQIOLD-----AFVEVQKALINE----- 169
Db 536 EAKKLAQRLQDAEHEVAVNSKASLENTKQRLQNEAEDLMIDVERSNTATCARMDKQ 595

QY 170 -----HGAQEQIN-----KIRDSYAKLLOH-----QNLK 194

```

596 RNFDPKVLAEWKHYETQAELEASQKESRSLSTEVFKVKNAYEESLDHLETLKRENKNLQ 655
 195 QKI-----KHVVKLKDNSQLKSEVSKLRQLAKKKOSQTKLOBELNKVLGK 242
 656 QEISDLTEQIAESAKHIELEKVKQTDQEKSELQAAL---EBAEGSLHEEGKILRIQ 711

RESULT 26
 MYSC_CHICK STANDARD; PRT; 1102 AA.
 AC P29616;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Myosin heavy chain, cardiac muscle isoform (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]_TaxID=9031;
 RP SEQUENCE OF 65-1102 FROM N.A.
 RC STRAIN=Broiler; TISSUE=Heart;
 RX MEDLINE=92130260; PubMed=1774788;
 RA Stewart A.F.R., Camoretti-Mercado B., Perlman D., Gupta M.,
 RA Jakovcic S., Zak R.;
 RT "Structural and phylogenetic analysis of the chicken ventricular
 RT myosin heavy chain rod.";
 RL J. Mol. Evol. 33:357-366(1991).
 RN [2]
 RP SEQUENCE OF 1-259.
 RC TISSUE=Heart;
 RX MEDLINE=93039740; PubMed=1418675;
 RA Watanabe B.;
 RT "Amino-acid sequence of the short subfragment-2 in adult chicken
 RT cardiac muscle myosin.";
 RL Biol. Chem. Hoppe-Seyler 373:1045-1054(1992).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -!- TISSUE SPECIFICITY: VENTRICULAR MUSCLE AND TRACES LEVELS IN THE
 CC ATRIUM. ALSO TRANSIENT EXPRESSION IN SKELETAL MUSCLE DURING FETAL
 CC DEVELOPMENT AND REGENERATION FOLLOWING FREEZE INJURY.
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -!- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
 CC WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
 CC CONSERVED.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X59552; CAA42130.1; -;
 DR HSSP; P03437; IHTM.
 DR InterPro; IPR002928; Myosin tail.
 DR Pfam; PF01576; Myosin_tail_1
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Multigene family.
 FT NON_TER 1
 FT DOMAIN <1 1102 COILED COIL (POTENTIAL).
 FT CONFLICT 65 A -> T (IN REF. 1).

FT CONFLICT 92 92 V -> A (IN REF. 1).
 FT CONFLICT 113 114 CS -> FAL (IN REF. 1).
 FT CONFLICT 135 135 H -> Q (IN REF. 1).
 FT CONFLICT 185 185 T -> R (IN REF. 1).
 SQ SEQUENCE 1102 AA; 128008 MW; 2293668D049825DC CRC64;
 Query Match 13.9%; Score 166.5; DB 1; Length 1102;
 Best Local Similarity 20.9%; Pred.No.0.12;
 Matches 71; Conservative 55; Mismatches 94; Indels 119; Gaps 11;
 QY 8 VOSLEDVTAQFESYKALTASBIEDLKLNSLSQEAQAKAGKNAEDV---QHOILATESG 63
 Db 373 LNLQVRQKLEKEKSELKWEYDDL---TANWEQTVKGAKEKLWGTYEDHLNETTK 428
 QY 64 NOEYVRMLDLQTKSA-----LKTEIKEITV-----SFLQKITDLQNLKQO 106
 Db 429 LDEMTRLNMDLTQTKLQSENGEFVQLEEKESLSQLSRGKTSFTQOIIEELRQLEEE 488
 QY 107 EED-----FRKQLEDEGRKAENKTAAELTEINKWELLY----- 142
 Db 489 TSKNALAHALQAARHDCDLLREQYEEBOEAKAELQALSKGNAEVAQWTKYETDAIQ 548
 QY 143 -EELYNKTKPFOIQI-----DAFEVEKQALLNEH-----GAAQEQ 177
 Db 549 TEELEDAKKLLARLQEAEEAEAEANAKCSLEKAKHRLQNEQEDMIDLEKANSAA 608
 QY 178 NKIRDSYAKLGHQNLKQK-----KQOSETKQ 233
 Db 609 DKKQGFQDKII--NDWKQYESSQAELEASQKESRSLSTELFKLNAYEETLDHLETLKR 666
 QY 206 ENSQLKSEVSKLRQLAK-----KKOSETKQ 233
 Db 667 ENKLQEEISDLTNOISGNKNLHEIEKVKQVEQEKSE 705

RESULT 27
 RA50_SULTO STANDARD; PRT; 879 AA.
 AC Q96TR5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN DNA OR S72108
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX MEDLINE=21456156; PubMed=11572479;
 RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000988; BAB67212.1; ALT_INIT.
DR HAMAP; MF 00449; -. 1.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003405; SMC_C.
DR InterPro; IPR003395; SMC_N.
DR Pfam; PF04423; Rad50_zn_hook; 1.
DR Pfam; PF02483; SMC_C; 1.
DR Pfam; PF02453; SMC_N; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR DNA repair; Hydrolase; ATP-binding; Coiled coil;
KW Complete proteome.
FT NP_BIND 30 37 ATP (BY SIMILARITY).
FT DOMAIN 172 731 COILED COIL (POTENTIAL).
SQ SEQUENCE 879 AA; 103487 MW; E9103B58914F20D7 CRC64;

Query Match 13.8%; Score 165.5; DB 1; Length 879;
Best Local Similarity 22.7%; Pred. No. 0.1;
Matches 77; Conservative 52; Mismatches 97; Indels 113; Gaps 14;

QY 8 VOSLEDVTAQFYSKALTASEIEDLKLENSLOEKAAGKAGKQAEVQHOILATESNQY 67
Db 186 LQSLKYKNESENQIKQKELENIKRELEDLNKKEKRYEDIVKLINEBEKKEKY 245
QY 68 VPM-----LIDLQTSAL 80
Db 246 VELISLLNKLKDDISELREBEVDENRLREBEKLEKLDLEKLIKEKITEAQNKIKL 305
QY 81 ---KETEIKETVTFPQKTDIQNLQKQ---DEDRKOLE--DEGRKAENKNTAET 132
Db 306 AQEKESLTIKIN-----LTDLEKLRKRELEEDYKKEIKELKELDELNKKFKLEINKIN 361
QY 133 EBINKRWLYEELYNKTKPFQIQDAFEVEKQAL-----LNEHGAAQ-----QLNKIR-- 181
Db 362 DRLSKSLKILSIESKISNRKISINIEEDLKLQKLNEDLNKNQREKLSQLEIKR 421
QY 182 -DSYAKLIGHQNL-----KQIKH--VVKLK--DE-NSOLKSEVSKLR 218
Db 422 IBEELNKLGLNQKGVNCPVCGRELSDDHKRQIQNEIEKLKELDELNKKFKLEINKIN 481
QY 219 -----QLAKKQSEFKL-----QBELNK 237
Db 482 GLISELNQIINKSKSEKDIARNLADYNNLLTQQQBELR 520

RESULT 28
CING_XENLA STANDARD; PRT; 1360 AA.
AC Q9PFD7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cingulin.
GN CGN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RA MEDLINE=20082893; PubMed=10613913;
RA Cordenonsi M., D'Atri F., Hammar E., Parry D.A.D., Kenrick-Jones J.,
RA Shore D., Citi S.;
RT "Cingulin contains globular and coiled-coil domains and interacts with
RT ZO-1, ZO-2, ZO-3 and myosin.";
RL J. Cell Biol. 147:1569-1581(1999).
RN [2]

FUNCTION IN TIGHT JUNCTION BIOGENESIS.
MEDLINE=97029068; PubMed=8875080;
Cardellini P., Davanzo G., Citi S.;
"tight junctions in early amphibian development: detection of
functional cingulin from the 2-cell stage and its localization at the
boundary of distinct membrane domains in dividing blastomeres in low
calcium.";
Dev. Dyn. 207:104-113(1996).
[3]
FUNCTION IN TIGHT JUNCTION BIOGENESIS.
MEDLINE=20400099; PubMed=10940624;
Fesenko I., Kurth T., Sheth B., Fleming T.P., Citi S., Hausen P.;
"tight junction biogenesis in the early Xenopus embryo.";
Mech. Dev. 96:51-65(2000).
[4]
INTERACTION WITH OCCLUDIN.
MEDLINE=99421641; PubMed=10491082;
Cordenonsi M., Turco F., D'Atri F., Hammar E., Martinucci G.,
Meggio F., Citi S.;
"Xenopus laevis occludin. Identification of in vitro phosphorylation
sites by protein kinase CK2 and association with cingulin.";
Eur. J. Biochem. 264:374-384(1999).
[5]
INTERACTION WITH F-ACTIN.
MEDLINE=21538627; PubMed=11682052;
D'Atri F., Citi S.;
"Cingulin interacts with F-actin in vitro.";
FEBS Lett. 507:21-24(2001).
[6]
INTERACTION WITH ZO-1.
MEDLINE=22140336; PubMed=12023291;
D'Atri F., Nadalutti F., Citi S.;
"Evidence for a functional interaction between cingulin and ZO-1 in
cultured cells.";
J. Biol. Chem. 277:27757-27764(2002).
-!- FUNCTION: Probably plays a role in the formation and regulation of
the tight junction (TJ) paracellular permeability barrier,
possibly by linking ZO proteins to the actomyosin cytoskeleton.
-!- SUBUNIT: Parallel homodimer (by similarity). Binds ZO-1 and ZO-2
in vivo, and ZO-3, myosin and occludin in vitro, possibly
directly. Acts as an F-actin bundling protein in vitro.
-!- TISSUE SPECIFICITY: Localized on the cytoplasmic face of tight
junctions of polarized epithelia and some endothelia.
-!- DEVELOPMENTAL STAGE: A maternally synthesized protein. Found in
the apical cortex in the fertilized egg, where it is associated
with cytoskeleton filaments, it is recruited to tight junctions
before ZO-1 and occludin. Nascent tight junctions are in place
by the two-cell stage.
-!- DOMAIN: Deletion of the ZO-1 interaction motif (ZIM) decreases but
does not abolish colocalization with ZO-1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

Query Match 13.8%; Score 165; DB 1; Length 1360;
 Best Local Similarity 23.1%; Pred. No. 0.17;
 Matches 64; Conservative 63; Mismatches 90; Indels 60; Gaps 10;

QY 10 SLEDVTAQESYKALTASIEDLKLENSIQEAKAGNAEDVQOILATESNQEVYR 69
 DB 1030 SLQLQTOQEYREKSRKEIGE-----AQQAKE-KTAAERHQF--NSRMQEEVQ 1078

QY 70 MLLDLQTSALKETEIKETIVSFLOKITDLOQLKQEBDFRKLQDEBGRKA----- 122
 DB 1079 KL-----KLALQELQVEKETVELDKQMSORLSQLEQDIJESKRVQDDRSQKVLEDKL 1133

QY 123 -----EKENTTALTEINKRWLLYEE-----LYNKTYPQ 153
 DB 1134 KMEAELEDEKNTVELLTDVNRSDQMEQORAEINQERSGDLCECDKISLERQNKELK 1193

QY 154 IOLDPEVEKQALLN-EHGAA-----QEOLNKIRDSYAKILG-HQNLKOKIKHV-VKLKD 205
 DB 1194 NLAASNEGQKQSVNVSHLEAKLOEIQERLQLEBEREKATILLSTNRKLERKLELNQLED 1253

QY 206 ENSQLKSEVSKLRQAKKQSBETKLQEBLNKVLGK 242
 DB 1254 ERLQVNDQXQLNLRVYKALKRQVDEAEIEERLEGUR 1290

RESULT 29
 MYH6_MOUSE
 ID MYH6_MOUSE STANDARD; PRT; 1938 AA.
 AC Q02566; Q64258; Q64738;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE MYH6 heavy chain, cardiac muscle alpha isoform (MYHC-alpha).
 GN MYH6 OR MYHCA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c, A/J, C57BL/6J, and DBA/2J;
 RX MEDLINE=92250040; PubMed=1577481;
 RA Quinn-Laquer B.K., Kennedy J.E., Wei S.J., Beisel K.W.;
 RT "Characterization of the allelic differences in the mouse cardiac
 alpha-myosin heavy chain coding sequence.";
 RL Genomics 13:176-188(1992).
 RN [2]
 RP SEQUENCE OF 1-67 FROM N.A.
 RC STRAIN=AKR.
 RX MEDLINE=91225025; PubMed=2026617;
 RA Gullick J., Subramaniam A., Neumann J., Robbins J.;
 RT "Isolation and characterization of the mouse cardiac myosin heavy
 chain genes.";
 RL J. Biol. Chem. 266:9180-9185(1991).
 CC - FUNCTION: MUSCLE CONTRACTION.
 CC - SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC - SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC - DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC - PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC - MISCELLANEOUS: THE CARDIAC ALPHA ISOFORM IS A 'FAST' ATPASE
 MYOSIN, WHILE THE BETA ISOFORM IS A 'SLOW' ATPASE.
 CC - SIMILARITY: Contains 1 myosin-like globular head domain.
 CC - SIMILARITY: Contains 1 IQ domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M76598; AAA37159.1; -
 DR EMBL; M76599; AAA37160.1; -
 DR EMBL; M76600; AAA37161.1; -
 DR EMBL; M76601; AAA37162.1; -
 DR EMBL; M62404; AAA37424.1; -
 DR PIR; I49464; I49464.
 DR HSSP; P08799; LMND.
 DR SWISS-2DPAGE; Q02566; MOUSE.
 DR MGD; MGI:97255; Myhca.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; Myosin_head.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS50096; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family;
 KW Calmodulin-binding; Polymorphism.
 FT DOMAIN 1 782 MYOSIN HEAD-LIKE.
 FT DOMAIN 783 812 IQ.
 FT DOMAIN 842 1938 COILED COIL (POTENTIAL).
 FT NP_BIND 178 185 ATP.
 FT DOMAIN 657 679 ACTIN-BINDING.
 FT DOMAIN 759 773 ACTIN-BINDING.
 FT DOMAIN 790 807 CALMODULIN-BINDING (BY SIMILARITY).
 FT DOMAIN 816 833 CALMODULIN-BINDING (BY SIMILARITY).
 FT MOD_RES 129 129 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 697 697 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 707 707 ALKYLATION (SH-2) (POTENTIAL).
 FT VARIANT 194 194 Y -> D.
 FT VARIANT 545 545 S -> A.
 FT VARIANT 838 838 I -> S.
 SQ SEQUENCE 1938 AA; 223564 MW; EAD789ADA68818FB CRC64;
 Query Match 13.6%; Score 162.5; DB 1; Length 1938;
 Best Local Similarity 21.1%; Pred. No. 0.33;
 Matches 76; Conservative 50; Mismatches 99; Indels 135; Gaps 12;

QY 8 VQSLDVTVAQESYKALTASIEDLKLENSIQEAKAGNAEDVQOILATESNQEVYR 67
 DB 1209 IDNLQVRQKLEKSEFKLEDDV-----TSNMEQITIKAKANLEKVSFTL---EQDAEY 1261

QY 68 -----VRMLLDLQTSALKETE-----IKEIT---VSFLOKITDLOQL 103
 DB 1262 RVKLEEAQRSLNDTTQPAKLOQTENGELARQLEKEALISQITRGKLSYTOQMEDLKKQL 1321

QY 104 KQEBED-----FRKQLEDEBGRKAENKTAEITBEINKRWLLY----- 142
 DB 1322 EEEGKAKVALAHALQSSRHDCDLLREQVEEEMEAKEQLRVLSKANSEVAQWRKYETDA 1381

QY 143 -----BELNKTYPQIQI-----DAFEVEKQALLNE-----HGAAQ 174
 DB 1382 IQRTEELEEAARKLAQRLOQDAEAEAVNAKCSSELEKTKHRLQNEIEDLMVDVERSNAAA 1441

QY 175 EQLNKRDSYAKILGHQNLK-----QKIKHVVKLK 204
 DB 1442 AALDKKGRNFDKILAEWKQKYEESQSELESQKARSLSLTFLPKLNAYEESLEHETFK 1501

QY 205 DENSOLKSEVS-----KLRQIAKKK-----QSETKLOEINKNVL 239
 Db 1502 RENKNLOEISDLTQLGEGGKNNVHELEKIRKQLEVEKLEQSALEAEASLEHEGKIL 1561
 RESULT 30
 MYH6 HUMAN
 ID MYH6 HUMAN STANDARD; PRT; 1939 AA.
 AC P13533; Q13943; Q14906; Q14907;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, cardiac muscle alpha isoform (MyHC-alpha).
 GN MYH6 OR MYHCA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92133665; PubMed=1776652;
 RA Matsuoka R., Beisel K.W., Furutani M., Arai S., Takao A.;
 RT "Complete sequence of human cardiac alpha-myosin heavy chain gene and
 RT amino acid comparison to other myosins based on structural and
 RT functional differences.";
 RL Am. J. Med. Genet. 41:537-547(1991).
 FN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94140346; PubMed=8307559;
 RA Epp T.A., Dixon I.M., Wang H.Y., Sole M.J., Liew C.-C.;
 RT "Structural organization of the human cardiac alpha-myosin heavy
 RT chain gene (MYH6).";
 RL Genomics 18:505-509(1993).
 FN [3]
 RP SEQUENCE OF 1-177 AND 1551-1939 FROM N.A.
 RX MEDLINE=89264452; PubMed=2726733;
 RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;
 RT "Characterization of human cardiac myosin heavy chain genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3504-3508(1989).
 FN [4]
 RP ERRATUM.
 RA Yamauchi-Takahara K., Sole M.J., Liew J., Ing D., Liew C.-C.;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7416-7417(1989).
 FN [5]
 RP SEQUENCE OF 1407-1939 FROM N.A.
 RX MEDLINE=88299163; PubMed=2969919;
 RA Kurabayashi M., Tsuchimotochi H., Komuro I., Takaku F., Yazaki Y.;
 RT "Molecular cloning and characterization of human cardiac alpha- and
 RT beta-form myosin heavy chain complementary DNA clones. Regulation of
 RT expression during development and pressure overload in human
 RT atrium.";
 RL J. Clin. Invest. 82:524-531(1988).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -!- MISCELLANEOUS: THE CARDIAC ALPHA ISOFORM IS A 'FAST' ATPASE
 CC MYOSIN, WHILE THE BETA ISOFORM IS A 'SLOW' ATPASE.
 CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
 CC -!- SIMILARITY: Contains 1 IQ domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D00943; BAA00791.1; -;
 DR EMBL; Z20656; CAA79675.1; -;
 DR EMBL; M25140; AAA60386.1; -;
 DR EMBL; M25162; AAA60386.1; JOINED.
 DR EMBL; M25142; AAA60387.1; -;
 DR EMBL; M25141; AAA60387.1; JOINED.
 DR EMBL; M21664; AAA6344.1; -;
 DR PIR; A46762; A46762.
 DR HSSP; P08799; 1MND.
 DR Genew; HGNC:7576; MYH6.
 DR MIM; 160710; -;
 DR GO; GO:0005859; C:muscle myosin; TAS.
 DR GO; GO:0003779; F:actin binding activity; NAS.
 DR GO; GO:0005524; F:ATP binding activity; NAS.
 DR GO; GO:0005516; F:calmodulin binding activity; NAS.
 DR GO; GO:0003776; F:muscle motor activity; NAS.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; myosin_head.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR Pfam; PF00612; IQ; 2.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS00096; IQ; 1.
 KW Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family;
 KW Calmodulin-binding.
 FT DOMAIN 1 782 MYOSIN HEAD-LIKE.
 FT DOMAIN 783 812 IQ.
 FT DOMAIN 842 1939 COILED COIL (POTENTIAL).
 FT NP_BIND 178 185 ATP.
 FT DOMAIN 657 679 ACTIN-BINDING.
 FT DOMAIN 759 773 ACTIN-BINDING.
 FT MOD_RES 129 129 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 697 697 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 707 707 ALKYLATION (SH-2) (POTENTIAL).
 FT CONFLICT 88 88 Q -> E (IN REF. 1).
 FT CONFLICT 574 574 Q -> P (IN REF. 1).
 FT CONFLICT 608 608 A -> G (IN REF. 1).
 FT CONFLICT 744 744 T -> A (IN REF. 1).
 FT CONFLICT 790 790 M -> I (IN REF. 1).
 FT CONFLICT 1014 1014 V -> A (IN REF. 1).
 FT CONFLICT 1021 1021 S -> T (IN REF. 1).
 FT CONFLICT 1101 1101 A -> V (IN REF. 1).
 FT CONFLICT 1290 1290 A -> S (IN REF. 1).
 FT CONFLICT 1373 1373 W -> C (IN REF. 1).
 FT CONFLICT 1533 1533 K -> N (IN REF. 5).
 FT CONFLICT 1540 1540 L -> M (IN REF. 5).
 FT CONFLICT 1577 1577 KL -> NV (IN REF. 5).
 FT CONFLICT 1706 1706 EQ -> DR (IN REF. 1).
 FT CONFLICT 1733 1733 E -> D (IN REF. 1).
 FT CONFLICT 1734 1734 A -> S (IN REF. 2).
 FT CONFLICT 1737 1737 T -> S (IN REF. 1).
 FT CONFLICT 1763 1763 D -> H (IN REF. 1).
 FT CONFLICT 1788 1788 M -> I (IN REF. 3).
 FT CONFLICT 1871 1871 D -> N (IN REF. 5).
 FT CONFLICT 1882 1882 R -> G (IN REF. 5).
 FT CONFLICT 1890 1890 Q -> R (IN REF. 5).
 FT CONFLICT 1933 1933 MISSING (IN REF. 5).
 FT SEQUENCE 1939 AA; 223689 MW; ECB87E7CE8768B6F CRC64;

Search completed: December 16, 2003, 06:08:03
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:07:28 ; Search time 34 Seconds
(without alignments)
1836.727 Million cell updates/sec

Title: US-09-978-309a-74
Perfect score: 1195
Sequence: 1 QEKYDSMVQSLDVTAFES.....KKQSETKLQBELNKVLGIK 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	917	76.7	836	11 Q8VDR2	Q8vdr2 mus musculu
2	905.5	75.8	476	11 Q920Z6	Q920z6 mus sp. hya
3	893	74.7	713	11 Q9WUF7	Q9wuf7 rattus norv
4	338	28.3	82	6 Q95JG7	Q95jg7 ovis aries
5	338	28.3	82	6 Q8SPM2	Q8spm2 bos taurus
6	209	17.5	436	4 Q9NS87	Q9ns87 homo sapien
7	209	17.5	1388	4 Q9NS87	Q9ns87 homo sapien
8	193.5	16.2	1944	13 Q9DGM5	Q9dgm5 gallus gall
9	193	16.2	1388	13 Q91785	Q91785 xenopus lae
10	191.5	16.0	1943	13 Q8UG72	Q8ug72 gallus gall
11	191	16.0	795	4 Q9H2G7	Q9h2g7 homo sapien
12	191	16.0	804	4 Q8TC31	Q8tc31 homo sapien
13	191	16.0	1410	4 Q14221	Q14221 homo sapien
14	191	16.0	1411	4 Q15075	Q15075 homo sapien
15	190	15.9	1690	5 Q44929	Q44929 drosophila
16	189.5	15.9	1941	13 Q9DGM4	Q9dgm4 gallus gall

17	189	15.8	1652	5 Q8INV8	Q8inv8 drosophila
18	189	15.8	1689	5 Q8MSD0	Q8msd0 drosophila
19	189	15.8	1690	5 Q9VJES	Q9vjes drosophila
20	187	15.6	826	13 Q9YHDS	Q9yhd5 rana catesb
21	185.5	15.5	1790	3 Q07380	Q07380 saccharomyc
22	183.5	15.4	1937	6 Q9TV62	Q9tv62 sus scrofa
23	182.5	15.3	1929	13 Q98TQ6	Q98tq6 notothenia
24	182.5	15.3	1935	13 Q902E5	Q902e5 brachydania
25	182.5	15.3	1940	6 Q9BE41	Q9be41 bos taurus
26	181.5	15.2	1598	11 Q922D2	Q922d2 mus musculu
27	180.5	15.1	1937	6 Q8MJV1	Q8mjv1 equus cabal
28	180.5	15.1	1938	6 Q9BB40	Q9bb40 bos taurus
29	179.5	15.0	1939	13 Q9PTY2	Q9pty2 gallus gall
30	179	15.0	1463	5 Q8GYZ0	Q8gyz0 strongyloce
31	178.5	14.9	692	5 Q967Z0	Q967z0 dermatophag
32	178.5	14.9	1930	13 Q9DGD5	Q9dgd5 pennahia ar
33	178.5	14.9	1939	6 Q9TV61	Q9tv61 sus scrofa
34	178	14.9	609	17 Q8TXA4	Q8txa4 methanopyru
35	177.5	14.9	1938	6 Q8MJV0	Q8mjv0 equus cabal
36	176.5	14.8	975	13 Q98TQ5	Q98tq5 notothenia
37	176.5	14.8	1939	6 Q9TV63	Q9tv63 sus scrofa
38	175.5	14.7	746	4 Q9HG6N	Q9h6n6 homo sapien
39	175.5	14.7	1945	6 Q97757	Q97757 felis silve
40	174.5	14.6	1936	13 Q90YF6	Q90yf6 paracirrhith
41	174.5	14.6	1937	13 Q8JTP5	Q8jtp5 oncorhynch
42	174	14.6	549	5 Q9XZV8	Q9xzv8 hydra atten
43	172.5	14.4	1940	13 Q8AX28	Q8ayx8 gallus gall
44	171	14.3	1320	11 Q9UK25	Q9jkz5 rattus norv
45	170.5	14.3	1119	13 P87344	P87344 theragra ch
46	170	14.2	712	11 Q8BR64	Q8br64 mus musculu
47	170	14.2	756	11 Q8CBR0	Q8cbr0 mus musculu
48	170	14.2	1388	11 P70336	P70336 mus musculu
49	169.5	14.2	1938	13 Q9IBD7	Q9ibd7 seriola dum
50	169	14.1	536	12 Q8US31	Q8js31 phthorimaea

ALIGNMENTS

RESULT 1

Q8VDR2	PRELIMINARY;	PRT;	836 AA.
AC Q8VDR2;			
DT 01-MAR-2002 (TRENBLrel. 20, Created)			
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)			
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)			
DE Similar to hyaluronan mediated motility receptor (RHAMM).			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Strausberg R.;			
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; BC021427; AAH21427.1; -			
KW Receptor.			
SQ SEQUENCE 836 AA; 96670 MW; B9BDDA22BEECA652 CRC64;			

Query Match 76.7%; Score 917; DB 11; Length 836;
Best Local Similarity 76.4%; Pred. No. 2.1e-40;
Matches 185; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY	1	QEKYDSMVQSLDVTAFESYKALTASIEDLKLENSLQEKAAKAGNAEDVQHOILAT 60	
DB	568	QEKYNDTAQSLRDVTAFLESYKSTLKEIEDLKLENTLQEKVAAEKSVEDVQOQILTA 627	
QY	61	ESSNOEYVRLDLQTKALKATEIKETVTSFLOKITTDLQNLKQOEDFRKQLEDEGR 120	
DB	628	ESTNOEYARMVDLQNRSLTKEEIKETTSFLEKITDLKNLQROQDSEDFRQLEBKGR 687	
QY	121	KAEKENTTAETEEINKWRLLYEELNYNTKPKQIQLDAFEVEKQALLNHEGAAQQLANKI 180	

```

Db 688 TAENKENVMTLWEINKWRLLYEELYEKTPFQQQLDAFEAEKQALLNEHGATQEQLNKI 747
QY 181 RDSYAKLLGHONLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSEYKLOEELNKVLG 240
Db 748 RDSYAQLLGHONLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSEYKLOEELNKVLG 807
QY 241 IK 242
Db 808 IR 809

RESULT 2
Q92026 PRELIMINARY; PRT; 476 AA.
AC Q92026;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE Hyaluronan receptor.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92299690; PubMed=1376732;
RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
RA Cripps V., Austen L., Nance D.M., Turley E.A.;
RT "Molecular cloning of a novel hyaluronan receptor that mediates tumor
RT cell motility.";
RL J. Cell Biol. 117:1343-1350(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92348516; PubMed=1639856;
RA Hardwick C., Hoare K., Owens R., Hohn H.P., Hook M., Moore D.,
RA Cripps V., Austen L., Nance D.M., Turley E.A.;
RT "Correction. Molecular cloning of a novel hyaluronan receptor that
RT mediates tumor cell motility.";
RL J. Cell Biol. 118:753-753(1992).
DR EMBL; S41029; AAA09809.1; -.
KW Receptor.
SQ SEQUENCE 476 AA; 55486 MW; 62DCA82045EE0CB2 CRC64;

Query Match 75.8%; Score 905.5; DB 11; Length 476;
Best Local Similarity 76.0%; Pred. No. 4.7e-40;
Matches 194; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 1 QKYSVMQSLDVTAFQESYKALTASEIEDLKLENSLSQEKAAKAGKNAEDVQHQILAT 60
Db 209 QKYNDAQSLRDVSAQSLSEYKSLTKEIEDLKLENLTLQEKVMAEKSVEDVQQQILTA 268
QY 61 ESSNQEVYRMLDLQTKSALKETEIKETVSPLOKITDLOLQKQOEEDFRKQLEDEGR 120
Db 269 ESTNQIYARWQLOLRSTLKEEIKETVSSFLKEITDLOLQKQOEEDFRKQLEDEGR 328
QY 121 KAEKENTTAELTTEINKWRLLYEELYEKTPFQIOLDAFEAEKQALLNEHGATQEQLNKI 180
Db 329 TAENKENVMTLWEINKWRLLYEELYEKTPFQIOLDAFEAEKQALLNEHGATQEQLNKI 387
QY 181 RDSYAKLLGHONLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSEYKLOEELNKVLG 240
Db 388 RDSYAQLLGHONLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSEYKLOEELNKVLG 447
QY 241 IK 242
Db 448 IR 449

RESULT 3
Q9WUF7 PRELIMINARY; PRT; 713 AA.
AC Q9WUF7;
DT 01-NOV-1999 (TREMBlrel. 12, Created)

```

```

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hyaluronan receptor RHAMM.
GN RHAMM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE DAWLEY;
RA Li X., Lynn B., Nagy J.I., Cattini P.A.;
RT "RHAMM cDNA from rat brain.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pheochromocytoma;
RA Lynn B.D., Li X., Cattini P.A., Nagy J.I.;
RT "Sequence, protein expression and erk association of the hyaladherin
RT RHAMM in PC12 cells.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF133037; RAD24473.1; -.
DR EMBL; AF336825; AAK21904.1; -.
KW Receptor.
SQ SEQUENCE 713 AA; 82395 MW; 2FE310D5759C6CB3 CRC64;

Query Match 74.7%; Score 893; DB 11; Length 713;
Best Local Similarity 74.0%; Pred. No. 3.1e-39;
Matches 179; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

QY 1 QKYSVMQSLDVTAFQESYKALTASEIEDLKLENSLSQEKAAKAGKNAEDVQHQILAT 60
Db 442 QKYSIDTAQTURDVTQALESYKSLTKEIEDLKLENLTLQEKVMAEKSVEDVQQQILTA 501
QY 61 ESSNQEVYRMLDLQTKSALKETEIKETVSPLOKITDLOLQKQOEEDFRKQLEDEGR 120
Db 502 ESTNQIYARWQLOLRSTLKEEIKETVSSFLKEITDLOLQKQOEEDFRKQLEDEGR 561
QY 121 KAEKENTTAELTTEINKWRLLYEELYEKTPFQIOLDAFEAEKQALLNEHGATQEQLNKI 180
Db 562 MTEKETAVTETLMEINKWRLLYEELYEKTPFQIOLDAFEAEKQALLNEHGATQEQLNKI 621
QY 181 RDSYAKLLGHONLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSEYKLOEELNKVLG 240
Db 622 RDSYAQLLGHONLKQIKHVVVKLKDENSQKSEVSKLRQOLAKKQSEYKLOEELNKVLG 681
QY 241 IK 242
Db 682 IR 683

RESULT 4
Q95JG7 PRELIMINARY; PRT; 82 AA.
AC Q95JG7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE Hyaluronan acid-mediated motility receptor (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Trachea;
RA Aoki T.A., Forteza R.M., Conner G.E.;
RT "Receptor for hyaluronan acid mediated motility in sheep (Ovis aries)
RT trachea.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF310973; AAK69578.1; -.
KW Receptor.

```

```

FT NON TER 1 1
FT NON TER 82 82
SQ SEQUENCE 82 AA; 9452 MW; 46B87555C0F1B66 CRC64;

Query Match
Best Local Similarity 28.3%; Score 338; DB 6; Length 82;
Matches 68; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 167 LNEHGAQQLNKRDSYAKLLGHQNLKQIKKHVVKLKDENSQKSEVSKLRCOLAKKKQ 226
DB 1 LNEHGAQQLNKRDSYAKLLGHQNLKQIKKHVVKLKDENSQKSEVSKLRCOLAKKKQ 60
QY 227 SETKLOEELNKVLGK 242
DB 61 SEAKLOEELNKVLGK 76

RESULT 5
Q8SPM2 PRELIMINARY; PRT; 82 AA.
AC Q8SPM2;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Receptor for hyaluronidic acid mediated motility (Fragment).
GN RHAMM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Schoenfelder M., Einspanier R.;
RT "Expression of HAS-system during oocyte maturation in the cow.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AT439694; CAD29126.1; -.
FT NON TER 1 1
FT NON TER 82 82
SQ SEQUENCE 82 AA; 9452 MW; 46B87555C0F1B66 CRC64;

Query Match
Best Local Similarity 28.3%; Score 338; DB 6; Length 82;
Matches 68; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 167 LNEHGAQQLNKRDSYAKLLGHQNLKQIKKHVVKLKDENSQKSEVSKLRCOLAKKKQ 226
DB 1 LNEHGAQQLNKRDSYAKLLGHQNLKQIKKHVVKLKDENSQKSEVSKLRCOLAKKKQ 60
QY 227 SETKLOEELNKVLGK 242
DB 61 SEAKLOEELNKVLGK 76

RESULT 6
Q96JX7 PRELIMINARY; PRT; 436 AA.
AC Q96JX7
ID Q96JX7
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ14910.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,

```

```

RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK027816; BAB55389.1; -.
KW Hypothetical protein.
SQ SEQUENCE 436 AA; 51433 MW; 3BE3ABEBD7A7821D CRC64;

Query Match
Best Local Similarity 17.5%; Score 209; DB 4; Length 436;
Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;

QY 1 QEKYDSMVQSLSDV-----TAQFESYKA-----LTASIEDLKLNSLSQEKA 43
DB 166 KNEYNFKMRQLEHVWDSAAEDPQSPKTPPHQTHLAKLLETQEQEIEDGRASKTSLEHLV 225
QY 44 AKAGKNAEDVQHILATESSNQEVYRMLLDQ---TKSALKETEIKELTVSFLOKITDQ 100
DB 226 TKLNEDREVKNABILRMKEQREMENLRLESQQLIEKNWLLQGQDDIK---RQKENSQ 282
QY 101 N-----QLKQOEDFRKQ-----LEDEGRKARKENTTAELTFEINKW 138
DB 283 NHPDNQQLKNEQESIKERLAKSIVBEMLKVKADLEEVQSALYNKMECLRMTEVERT 342
QY 139 RLLVEELYNKTPFQIOLDFAFEVQKALLNEHGAQQLNKRDSYAKLLGHQNLKQIK 198
DB 343 QTLSEKAFQEKELSKLEENYERERTSQEMEMLRKQVECLAENGLVGHQNLHOKIQ 402
QY 199 HVVKLKDENSQKSEVSKLRCQ---LAKKQSET 229
DB 403 YVRLKENVRLAEETEKLRANVFLKPKKRS 436

RESULT 7
Q9NS87 PRELIMINARY; PRT; 1388 AA.
ID Q9NS87
AC Q9NS87;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Kinesin-like protein 2.
GN HKLP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20435852; PubMed=10878014;
RA Sueishi M., Takagi M., Yoneda Y.;
RT "The Forkhead-associated Domain of Ki-67 Antigen Interacts with the
RT Novel Kinesin-like Protein HKlp2.";
RL J. Biol. Chem. 275:28888-28892(2000).
DR EMBL: AB035898; BAB03309.1; -.
DR HSSP; P17119; 3KAR.
DR Genew; HGNC:17273; KNSL7.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 1388 AA; 160160 MW; E127EB4B991CA83A CRC64;

Query Match
Best Local Similarity 17.5%; Score 209; DB 4; Length 1388;
Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;

QY 1 QEKYDSMVQSLSDV-----TAQFESYKA-----LTASIEDLKLNSLSQEKA 43
DB 1118 KNEYNFKMRQLEHVWDSAAEDPQSPKTPPHQTHLAKLLETQEQEIEDGRASKTSLEHLV 1177

```

QY 44 AKAGNAEDYHQIATLATESNQEVYRMLDLQ-----TKSALKETEIKETITVSFLOKITDLQ 100
 Db 1178 TKLNEDREVKNAEILRMKEQLREMLNRLSQQLEIKENWLLQGLDDIK---RQKNSDQ 1234
 QY 101 N-----QKQOEEDFRKQ-----LEDEGRKAEKENTTAELTEERINKW 138
 Db 1235 NHPDNQQLKNEQESIKERLAKSKIVEMLKKADLBEVQSALYNKNECRLMFDVERT 1294
 QY 139 RLIIYEELYNKTKPQIOLDAPFVEVEKQALLNHHGAQBLNKIRDSYAKLLGHQMLKOKIK 198
 Db 1295 QTLSEKAFQEQELRSKLEEMYEERERTSQEMEMLRKQVECLABENGKLVGHQNLHQIKQ 1354
 QY 199 HVVKLKDENSOLKSEVSKLRQ---LAKKXQSET 229
 Db 1355 YVRLKKNVRLAETSKLRAENVFLKRRSES 1388

RESULT 8
 Q9DGM5 PRELIMINARY; PRT; 1944 AA.
 ID Q9DGM5
 AC Q9DGM5
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Past myosin heavy chain isoform 2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Q., Bandman E.;
 RT "Seven skeletal myosin heavy chain genes (MyHC) are organized as a
 RT multigene complex in the chicken genome."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF272033; AAF99314.1; -
 DR HSSP; P13538; 2MYS.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; myosin_head.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR Pfam; PF00612; IQ; 1.
 DR Pfam; PF00663; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PS50096; IQ; 1.
 SQ SEQUENCE 1944 AA; 223211 MW; 154F438220072D68 CRC64;

Query Match 16.2%; Score 193.5; DB 13; Length 1944;
 Best Local Similarity 21.3%; Pred. No. 0.025;
 Matches 76; Conservative 53; Mismatches 106; Indels 121; Gaps 9;

QY 8 VOSLEDVTAQFESYKALTASEIEDLKLSSLOEKAKAGNAEDVHQIILATESNQBY 67
 Db 1215 IDNLQVQKQLEKSELKQVEIDDLASNMESVSKANLEKMKRTLEDQLSKTKSEEBH 1274
 QY 68 VRLMLDLQTKSALKETEIKETITVSFLOKITDLQKQO-----VSLQKITDIQNLKQO 106
 Db 1275 QRMINDLSTQRLALQTESYGRQVEEKQALLISQISRGQAFQTQIEELKRLHLEEIKAK 1334
 QY 107 -----EED---FRKQLEDEGRKAEKENTTAELTEERINKWRLY-----BEL 145
 Db 1335 NALAHALQSGARHDCDLLRQYEEQKAGELQALSKANSEVAQWTKYETDAIQTTEL 1394
 QY 146 YNKTFFQIOLD-----AFVEKQALLNE-----HGAQEQEQLNKIR 181
 Db 1395 EEAKKLAQRLQDABEHBVAVNAKASLEKTKQRLQNEVEDLMIDVERANAACARLDKQ 1454

QY 182 DSYAKLLGHQNLK-----QKIKHVVKLKDENSOLK 211
 Db 1455 KNFDKILAEWKQKYETQAELEASQKESRSLSTELPFMKNAVEESLDHLETIKRENKQLQ 1514
 QY 212 SEVSKLRQALAK-----KKQ-----SETKLOFELNKVLGIK 242
 Db 1515 QETSDLTEQIABGGKAIHELEKVKQIQEQFSKIQAELEAEASLEHEEGKILRLQ 1570

RESULT 9
 Q91785 PRELIMINARY; PRT; 1388 AA.
 ID Q91785
 AC Q91785
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE KLP2 protein.
 GN KLP2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Boleti H., Karsenti E., Vernos I.;
 RT "Xklp2, a new Xenopus centrosomal kinesin-like protein required for
 RT centrosome separation during mitosis."
 RL Cell 0:0-0(0).
 DR EMBL; X94082; CAA63826.1; -
 DR HSSP; P17119; 3KAR.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
 DR KW ATP-binding; Coiled coil; Microtubules; Motor protein.
 SQ SEQUENCE 1388 AA; 159141 MW; 1F8925558B9AE28E CRC64;

Query Match 16.2%; Score 193; DB 13; Length 1388;
 Best Local Similarity 22.9%; Pred. No. 0.019;
 Matches 71; Conservative 51; Mismatches 92; Indels 96; Gaps 7;

QY 9 QSLDVDTAQFESYKALTASEIEDLKLSSLOEKAKAGNA 50
 Db 1079 BELEKLEAFNKQEAALLTHEKEIVKEQQLSELNQVKLMTDLISREQEKIRASSNS 1138
 QY 51 -----BDVQHQIILATE-----SSNOYVYRMLDLQTK 77
 Db 1139 SSPVVLPTPTPEGNPYDSEIANLQKNTNLILVSEINBERTSKNBEIIRL-----K 1192
 QY 78 SALKETEIKETITVSFLOKI-TDLQNLK-----QOEEDFRKQLEDEGRKA 122
 Db 1193 MOLCETENWRLIEIQNLQCKMKELSKOLENCNNMKDSNDQKPSNDQDLKREIEKESVSRM 1252
 QY 123 EKENTTA-----ELTEENKRLLYEELYNKTKPQIOLD 158
 Db 1253 EKGKATEHILKQAELEBTRNLCTKHSNLSELSKEIERTSLEAKATEKEEIRSIILEG 1312
 QY 159 FEVEKQALLNHHGAQEQELNKIRDSYAKLLGHQMLKOKIKHVVKLKDENSOLKSEVSKLR 218
 Db 1313 KYEETKLSHEDMLRKQVFLABENGKILGHQNPQKIQIYLVKLVKKNKLLSEAEKLR 1372
 QY 219 COLAKKKQSE 228
 Db 1373 IENLFLKESK 1382

RESULT 10
 Q8JG72 PRELIMINARY; PRT; 1943 AA.
 ID Q8JG72

QY 86 -----KEITVSFLQKITDLOLQKQBEDPRKQLEDEBGRKARKENTTAELTEINWK 138
 Db 521 NKLSLENKELT-----QETSDMTLELKNQOEDINNNKQBERMLKQIENIQETETQLRNEL 576
 QY 139 RLLYBELNNTKPFQIOLDAFE-----VEKQALLNHHGAQBSOL 177
 Db 577 EYVREELQKRDVKLDSENCNNLRKOVENKNKVIIEELQOENKALKKKGTAEKQL 636
 QY 178 N-----KIRDSYA-----KLGHONLKQIKHVVVKLDKENSOLK 211
 Db 637 NVYEIKVKNLELESASAKQFGEITDYOKETEDKKISEENLEEVKAKVTADEAVKLQ 696
 QY 212 SEVSKLRQCO-----LAKKQSKSETKQLEELNKLVI 241
 Db 697 KEIDK-RCQHKTAENVALMEKHQYDKIIEERDSLGL 734
 RESULT 13
 ID Q14221 PRELIMINARY; PRT; 1410 AA.
 AC Q14221;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)
 DE Endosome-associated protein.
 GN EEAL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95286647; PubMed=7768953;
 RA Wu F.T., Callaghan J.M., Steele-Mortimer O., Stenmark H., Parton R.G.,
 RA Campbell P.L., McCluskey J., Yeo J.P., Tock E.P., Toh B.H.;
 RT "EEAL, an early endosome-associated protein. EEAL is a conserved
 RT alpha-helical peripheral membrane protein flanked by cysteine
 RT 'fingers' and contains a calmodulin-binding IQ motif.";
 RL J. Biol. Chem. 270:13503-13511(1995).
 DR EMBL; I40157; AAA79121.1; --
 DR Genew; HGNC:3185; EEAL.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR000306; Znf_FYVE.
 DR Pfam; PF01363; FYVE; 1.
 DR SMART; SM00064; FYVE; 1.
 DR SMART; SM00355; Znf_C2H2; 1.
 DR PROSITE; PS0178; ZF FYVE; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 1410 AA; 162496 MW; DF0F9464D70A8AED CRC64;

Query Match 16.0%; Score 191; DB 4; Length 1410;
 Best Local Similarity 27.7%; Pred. No. 0.025;
 Matches 74; Conservative 50; Mismatches 69; Indels 74; Gaps 12;
 QY 1 QEKYDSMVQSLDVTAFESYKALTASIEDLKLENSLOEKAAGKNAEDVQHQ----56
 Db 715 KEKYLSEQKTELEGQIK-----KLEADSLEVSKASKE-QALQDLQOORQLN 760
 QY 57 -----ILATESNQ-----EYVRML-LDLOTKSALKETIKEITVSFLQKITDLOLQKQ 106
 Db 761 TDLELRATELSKOLEMEKEIVSSTRLDLQKKS-----EALSIKQKLT-----KQE 806
 QY 107 EEDFRKOLEDEBGRKAEKENTTAELTEINRWLLYEELNKKTPFQIOLDAFEVSKQAL 166
 Db 807 EE---KQILKQDFETLSQET-----KIQHEELNNRIQTTVTLEQKVKMEKAL 851
 QY 167 LNEHGAQAEQLNKIRDSY-----AKLLGHONLKQIKHVVVKLDKENSOLK 211
 Db 852 MTELSTVKDKLSKVSLSKNSEFEKENQKGAAILDLKCTCKELKHQLOVQMENT-LK 910

QY 212 SEVSKLRQCLAKKQSKSETKQLEELNKV 238
 Db 911 -EQELKKSLEKEKEASHQKLKELNSM 936
 RESULT 14
 ID Q15075 PRELIMINARY; PRT; 1411 AA.
 AC Q15075;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)
 DE Endosomal protein.
 GN P162.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seelig H.P.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X78998; CAA55632.1; --
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR000306; Znf_FYVE.
 DR Pfam; PF01363; FYVE; 1.
 DR SMART; SM00064; FYVE; 1.
 DR SMART; SM00355; Znf_C2H2; 1.
 DR PROSITE; PS0178; ZF FYVE; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 1411 AA; 162465 MW; C3B17777FE34B6BD CRC64;
 Query Match 16.0%; Score 191; DB 4; Length 1411;
 Best Local Similarity 27.7%; Pred. No. 0.025;
 Matches 74; Conservative 50; Mismatches 69; Indels 74; Gaps 12;
 QY 1 QEKYDSMVQSLDVTAFESYKALTASIEDLKLENSLOEKAAGKNAEDVQHQ----56
 Db 715 KEKYLSEQKTELEGQIK-----KLEADSLEVSKASKE-QALQDLQOORQLN 760
 QY 57 -----ILATESNQ-----EYVRML-LDLOTKSALKETIKEITVSFLQKITDLOLQKQ 106
 Db 761 TDLELRATELSKOLEMEKEIVSSTRLDLQKKS-----EALSIKQKLT-----KQE 806
 QY 107 EEDFRKOLEDEBGRKAEKENTTAELTEINRWLLYEELNKKTPFQIOLDAFEVSKQAL 166
 Db 807 EE---KQILKQDFETLSQET-----KIQHEELNNRIQTTVTLEQKVKMEKAL 851
 QY 167 LNEHGAQAEQLNKIRDSY-----AKLLGHONLKQIKHVVVKLDKENSOLK 211
 Db 852 MTELSTVKDKLSKVSLSKNSEFEKENQKGAAILDLKCTCKELKHQLOVQMENT-LK 910
 QY 212 SEVSKLRQCLAKKQSKSETKQLEELNKV 238
 Db 911 -EQELKKSLEKEKEASHQKLKELNSM 936
 RESULT 15
 ID O44929 PRELIMINARY; PRT; 1690 AA.
 AC O44929;
 DT 01-JUN-1998 (TREMELrel. 06, Created)
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
 DE Microtubule binding protein D-CLIP-190.
 GN CLIP-190 OR CG5020.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;


```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STEAIN=OREGON R;
RA  Lantz V.A., Miller K.G.;
RL  J. Cell Biol. 0:0-0(1998).
RD  EMBL; AF041382; AAB96783.1;
DR  FlyBase; FBgn020503; CLIP-190.
DR  InterPro; IPR000938; CAP-GLY.
DR  Pfam; PF01302; CAP_GLY_2.
DR  PROSITE; PS0245; CAP_GLY_2; 2.
SQ  SEQUENCE 1690 AA; 189103 MW; BE4F48FD15F17A7C CRC64;

Query Match      15.9%; Score 190; DB 5; Length 1690;
Best Local Similarity 27.1%; Pred. No. 0.033;
Matches 75; Conservative 53; Mismatches 107; Indels 42; Gaps 11;

Qy  2 EKYDSMVQSLSDVTAQFESYKALTASEIEDLKLENSLSLOEKAA---KAGKNAEDVQHIL 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  629 EQIRELNQQLDEVITTLNVQKADSSALDMLRLQKEGTEEKSTLLEKTEKELVQKEQAA 688
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  59 ATESSNGEYVEMLDLQTKSALKETEIKEITVVSFLQKI-----TDLQNLKQOE-EDFR 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  689 KTLQDKQLEKQISDLK-QLAEQKELVREKTEINAINIQLEKESIEQQLALKQNELEDFQ 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  112 KQLEDEGR-----KAEKENTTTAELTEINKWRLYEELYNKT---KPFQIQLDAFEVEKQ 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  748 KQOSEVHLQEIKAQNTQKDLVSGESLKKLQQOLEEKLGHKLEKLOALEBELKKEKE 807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  165 ALLNEHGAQQL-NKIRDSVAKL-----LGHONLKQKI-----KHVVKLKDENSOLK 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  808 TIIRKEQEQLQQLQSKAESLSALKVQVQLEQLQQQAAASGEGSKTVAKLHDEISOLK 867
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  212 SEVSKLRCLQ-----AKKQSEK---KLOEELNK 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  869 SQAEETOSLKSSTENLEAKSKQLEANGSLEEEANK 904
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 16
Q9DGM4      PRELIMINARY;      PRT; 1941 AA.
AC  Q9DGM4;
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Fast myosin heavy chain isoform 3.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.;
RA  Zhang Q., Bandman E.;
RT  "Seven skeletal myosin heavy chain genes (MyHC) are organized as a
RL  multigene complex in the chicken genome."
RL  Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
DR  EMBL; AF272034; AAF99315.1;
DR  HSP; P13538; 2MYS.
DR  InterPro; IPR000048; IQ region.
DR  InterPro; IPR001609; myosin head.
DR  InterPro; IPR004009; Myosin N.
DR  InterPro; IPR002928; Myosin_tail.
DR  Pfam; PF00612; IQ; 2.
DR  Pfam; PF00063; myosin_head; 1.
DR  Pfam; PF02736; Myosin_N; 1.
DR  Pfam; PF01576; Myosin_Tail; 1.
DR  PRINTS; PR00193; MYOSINHEAVY.
DR  ProDom; PD000355; myosin_head; 1.
DR  SMART; SM00015; IQ; 1.
DR  SMART; SM00242; MYSC; 1.
DR  PROSITE; PS00096; IQ; 1.
SQ  SEQUENCE 1941 AA; 223162 MW; 9C8597C1CCF1DEBC CRC64;

```

```

Query Match      15.9%; Score 189.5; DB 13; Length 1941;
Best Local Similarity 21.1%; Pred. No. 0.041;
Matches 75; Conservative 53; Mismatches 107; Indels 121; Gaps 9;

Qy  8 VQSLSDVTAQFESYKALTASEIEDLKLENSLSLOEKAAKAGKNAEDVQHILATSSNQEY 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  1212 IDNLQRVQKLEKEKSELKMEIDDLASNMESVSKAKANLEKCMCTLEDQLSEIKTKEEH 1271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  68 VRMLDLQTKSALKETEIKEIT-----VSFLQKITDNLQNLKQ----- 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  1272 QRMINDLNTQARLQTEAGEYSRQVEEKDALISLSRGKQAFQOQIBELKHLHEEIKAK 1331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  107 -----EED---FRKLEDEEGSKAEKENTTTAELTEINKWRLY-----EEL 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  1332 NALAHALQSRHCDLLRQVEEEOEAKGLORALSANSVAQWRKYETDAIQTEEL 1391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  146 YNKTQPFQIQLD-----AFVEKQALLNE-----HGAAQQLNKIR 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  1392 EAQKKLAQRLQDAEHHVAVNAKASLEKTKQRLQNEVEDLMDVERANAACAALDKQ 1451
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  182 DSAKLIGHQNLK-----KQ-----SETKLOEELNKVLGK 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  1452 KNFDKILAEWKQYETQAELEASQKESLSLSTELFTKMNAYEESLDHLQTLKGNKQLQ 1511
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy  212 SEVSKLRCLQAK-----KQ-----SETKLOEELNKVLGK 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  1512 QEISDLTEQIAGGKAHELEKVKQLEQKSEIQAALEAEASLEHEGKILRLQ 1567
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 17
Q8INY8      PRELIMINARY;      PRT; 1652 AA.
AC  Q8INY8;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  CG5020-PC.
GN  CLIP-190.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20196006; PubMed=107311132;
RA  Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers J.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA  Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Flesler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Mobarry B., Murphy J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
RA Banon J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleeb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phanavanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochink S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnikier S.E.,
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase; (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003655; AAN10987.1; -;
SQ SEQUENCE 1652 AA; 186351 MW; B47CB31029196839 CRC64;
Query Match 15.8%; Score 189; DB 5; Length 1652;
Best Local Similarity 27.1%; Pred. No. 0.037;
Matches 75; Conservative 54; Mismatches 106; Indels 42; Gaps 11;
QY 2 EKYSVMOSLEDVTAQFESYKALTASEIEDLKLENSLSLOEKAA---KAGKNAEDVQHQL 58
Db EQIRELNQQLDEVTTQLNVQKADSSALDDMLRLQEGTEKSTLLEKTEKELVQSKQAA 650
QY 59 ATESNQEVVRMLLDLQTSKALKEITEIKTIVSFLQKI-----TDLQNLKQOE-EDFR 111
Db KTLNDKQLEKQISDLK-QLAEQKLVREMTENAINQIQLEKESIEQQLAKQNELEDFQ 709
QY 112 KOLEDEGR---KAEKENTTAEITELINKWRLLYBELYNKT---KPFQIQDADAEVEKQ 164
Db KKQSESEVHLQEIKAQNTQKDFELVESGESLKKLQQLQEGTKLGHEKLAALAEELKKEKE 769
QY 165 ALLNEHGAQEQ-LNKIRDSYAKL-----LGHQNLKOKI-----KHVVKLKDNSOLK 211
Db TIIKEKEQELQLOSKASRESALKVQVQLEQLQQAASGEGSKTVAKLHDEISOLK 829
QY 212 SEVSKLRQQL-----AKKQSET---KLQELNK 237
Db SOAETQSELKSTQSNLEAKSKQLEAANGSLEEAACK 866
RESULT 19
QYVJES PRELIMINARY; PRT; 1690 AA.
ID Q9VJES5
AC Q9VJES5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CLIP-190 protein.
GN CLIP-190 OR CG5020.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RESULT 18
Q8MSDO PRELIMINARY; PRT; 1689 AA.
ID Q8MSDO5
AC Q8MSDO5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE LD05834p.
GN CLIP-190 OR CG5020.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleeb J., Paragas V., Park S.,
RA Patel S., Phanavanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118896; AAM50756.1; -;
DR FlyBase; FBGN0020503; CLIP-190.
DR InterPro; IPR000938; CAP-GLY.
DR Pfam; PFO1302; CAP_GLY; 2.
DR PROSITE; PS50245; CAP_GLY; 2.
SQ SEQUENCE 1689 AA; 186936 MW; 410773DA3899E2E1 CRC64;
Query Match 15.8%; Score 189; DB 5; Length 1689;
Best Local Similarity 27.1%; Pred. No. 0.038;
Matches 75; Conservative 54; Mismatches 106; Indels 42; Gaps 11;
QY 2 EKYSVMOSLEDVTAQFESYKALTASEIEDLKLENSLSLOEKAA---KAGKNAEDVQHQL 58
Db EQIRELNQQLDEVTTQLNVQKADSSALDDMLRLQEGTEKSTLLEKTEKELVQSKQAA 687
QY 59 ATESNQEVVRMLLDLQTSKALKEITEIKTIVSFLQKI-----TDLQNLKQOE-EDFR 111
Db KTLNDKQLEKQISDLK-QLAEQKLVREMTENAINQIQLEKESIEQQLAKQNELEDFQ 746
QY 112 KOLEDEGR---KAEKENTTAEITELINKWRLLYBELYNKT---KPFQIQDADAEVEKQ 164
Db KKQSESEVHLQEIKAQNTQKDFELVESGESLKKLQQLQEGTKLGHEKLAALAEELKKEKE 806
QY 165 ALLNEHGAQEQ-LNKIRDSYAKL-----LGHQNLKOKI-----KHVVKLKDNSOLK 211
Db TIIKEKEQELQLOSKASRESALKVQVQLEQLQQAASGEGSKTVAKLHDEISOLK 866
QY 212 SEVSKLRQQL-----AKKQSET---KLQELNK 237
Db SOAETQSELKSTQSNLEAKSKQLEAANGSLEEAACK 903
RESULT 19
QYVJES PRELIMINARY; PRT; 1690 AA.
ID Q9VJES5
AC Q9VJES5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CLIP-190 protein.
GN CLIP-190 OR CG5020.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]


```

DR EMBL; Z74105; CAA98620.1; -.
DR SGD; S0002216; US01.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR006955; Usol_p115_C.
DR InterPro; IPR006953; Usol_p115_head.
DR Pfam; PF04871; Usol_p115_C_1.
DR Pfam; PF04869; Usol_p115_head; 1.
SQ SEQUENCE 1790 AA; 206450 MW; 90062544F55A52EE CRC64;

Query Match      15.5%; Score 185.5; DB 3; Length 1790;
Best Local Similarity 25.3%; Pred. No. 0.061;
Matches 72; Conservative 53; Mismatches 101; Indels 59; Gaps 10;

Qy 5 DSMVQSLDVTAQFESYKALTAS-----FIEDKLKNSLSLOEKAAGKAGNADVQHQI 57
Db 1148 ESLEKEHEDLAOLKKYBEQIANKERQYNBEISQLNDEITTOGENSIKKKNDELEGEV 1207

Qy 58 LATSSNCEYVRM-----LLDLOTKSALKETEITV-----SFLOKITDLOLQOL 103
Db 1208 KAMKSTSEQNLKSEIDALNLQIKELKKNETNEASLLSIKSEVETVKIKELQDEC 1267

Qy 104 XQOEEDFRKQLEDEGRKAENKTTAELTEBINKWRLLYBELYNKTKPFQIQLDAF----- 159
Db 1268 NFEKEV-SELEDKLCASEDKNSKYLEQKSEK---IKESLDAKTTTELKIQLEKITNLS 1323

Qy 160 -----EVE-----KQALLNEHGAQEOINKIR-----DSYAKLIGH-----QNL 193
Db 1324 KAKEKSELSRLKTKTSEERNAEQLKLEKIQIKNQAFKPKLLNKGSTITQEY 1383

Qy 194 KQIK----HVVKLDNSQLKSEVSKLRCLAKKKQSETKLOEE 234
Db 1384 SEKINTLEDELIRLQENELKAKEIDNTRSELEKYSLNDELLEE 1428

RESULT 22
Q9TV62 PRELIMINARY; PRT; 1937 AA.
AC Q9TV62;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Myosin heavy chain 2b.
OS Sus scrofa (pig)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactylia; Suina; Suidae; Sus.
CX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Landrace; TISSUE=Skeletal muscle;
RA Chikuni K., Tanabe R., Muroya S., Nakajima I.;
RT "Differences in molecular structure among the porcine myosin heavy
chain-2a, -2x, and -2b isoforms.";
RL Meat Sci. 57:311-317(2001).
DR EMBL; AB025261; BAA82145.1; -.
DR HSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 1937 AA; 223235 MW; BCC114C6824E0426 CRC64;

Query Match      15.4%; Score 183.5; DB 6; Length 1937;
Best Local Similarity 21.2%; Pred. No. 0.083;
Matches 77; Conservative 59; Mismatches 92; Indels 135; Gaps 11;

```

```

Qy 8 VQSLDVTAQFESYKALTASIEDLKLSLSLOEKAAGKAGNAE-----DVQHQILATESS 63
Db 1209 IDNLQVYKQKLEKEKSELKMEIDDL-----ASNMTYVSKAGNLEKMKRTLEDQLSEVTK 1264

Qy 64 NOEYVRMLLLOTKSALKETEI-----KRTVS-----FLQKITDLOLQOLQKQ 106
Db 1265 EEEHQRLNELSAQKARLQTESGEFSPQLDEKALVSQSRGKAQFTQQLIEELKRLQEE 1324

Qy 107 EED-----FRKQLEDEGRKAKENKTTAELTEBINKWRLLY----- 142
Db 1325 TKAKSALAHAVQSSRHDCDLLRQYEEQEAELQKAMSKANSEVAQWRTKYETDAIQ 1384

Qy 143 -EELYNKTKPFQIQLD-----AFVEKQALLNE-----HGAAGQOL 177
Db 1385 TEELEAKKLAQLODAREHVAVNAKCASEKTKORLQNEVEDLMDLVERSNAACAAL 1444

Qy 178 NKIRDSYAKLIG-----H 190
Db 1445 DKKQNFDKILAEWKHKYEEQTAELEASQKESRSLSTELFKVKNAYEESLDQLETKREN 1504

Qy 191 QNLKQKI-----KHVVKLDNSQLKSEVSKLRCLAKKKQSETKLQBELKVL 239
Db 1505 KNLQOEISDLTEQIAEGGKIHELEKVKQIOEKSELOAAL-----EBAEASLEHEGKIL 1561

Qy 240 GIK 242
Db 1562 RIQ 1564

RESULT 23
Q98TQ6 PRELIMINARY; PRT; 1929 AA.
AC Q98TQ6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myosin heavy chain.
GN MYOHC-A1.
OS Notothenia coriiceps (black rockcod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
OC Notothenioidae; Nototheniidae; Notothenia.
CX NCBI_TaxID=8208;
RN [1]
RP SEQUENCE FROM N.A.
RA Gauvry L., Ennion S., Ettelaie C., Goldspink G.;
RT "Characterisation of red and white muscle myosin heavy chain gene
coding sequences from Antarctic and tropical fish.";
RL Comp. Biochem. Physiol. 127:575-588(2000).
DR EMBL; AJ243767; CAC27776.1; -.
DR HSP; P13538; 2MYS.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 1929 AA; 221167 MW; 20A78CBA19DE851B CRC64;

Query Match      15.3%; Score 182.5; DB 13; Length 1929;
Best Local Similarity 21.2%; Pred. No. 0.094;
Matches 77; Conservative 65; Mismatches 86; Indels 135; Gaps 12;

Qy 8 VQSLDVTAQFESYKALTASIEDLKLSLSLOEKAAGKAGNAE-----DVQHQILATESS 63

```



```

QY 68 VRMLDLDTQTSALKETEI-----KEITVS-----FLOKITDLOQLKQO---- 106
DB 1272 ORLINDLTQRLQTESGPFQRQDEKALVSQLSRGKQAFQTOIEELKQLESEIKAK 1331
QY 107 -----BED-----FRKQLEDEGRKAEKENTTAELTEELINKWRLLY-----EEL 145
DB 1332 NALAHGLOSARHDCDLLREQYEEBESKAELOKALSANTVAQWRTKYETDAIORTEL 1391
QY 146 YNKTQFQIQLDA-----FEVEKQALLNE-----HGAAQEOQLNKIR 181
DB 1392 EAKKKLAQRLQAAEHEVAVNAKASLEKTKORLQNEVEDLMLDVERTNAACAALDKKQ 1451
QY 182 DSYAKLLG-----HQNL-----KQKIKHVVKLKDENSOLK 211
DB 1452 RNFDKILAEWKQYETHAELEAAQKARSIGTGLFKMKNAYEESLDQLETLKREKNLQ 1511
QY 212 SVSKLRCOLAK-----KKQ-----SETKLOEELNKVLGK 242
DB 1512 QEISDLTEQIAGGKMHLEKIKKQVEQKSEIQAALEAEASLEHEEGKILRIQ 1567

RESULT 26
Q922D2 PRELIMINARY; PRT; 1598 AA.
AC Q922D2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to myosin, heavy polypeptide 2, skeletal muscle, adult.
GN MYH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008538; AAH08538.1; -.
DR MGD; MGI:1339710; Myh2.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; myosin_N.
DR InterPro; IPR002928; Myosin tail.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin head; 1.
DR Pfam; PF02736; myosin_N; 1.
DR Pfam; PF01576; Myosin tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; P850096; IQ; 1.
SQ SEQUENCE 1598 AA; 183083 MW; D896GB343464A479 CRC64;

Query Match 15.2%; Score 181.5; DB 11; Length 1598;
Best Local Similarity 22.8%; Pred. No. 0.088;
Matches 82; Conservative 51; Mismatches 98; Indels 129; Gaps 12;

QY 8 VOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAE---DVQHQILATES 63
DB 1214 IDNLQVQKLEKEKSEMKMEIDDL-----ASNVTVSQKAGNLEKVCRTLEDQVSELKSK 1269
QY 64 NOEYVRLMLDLQTSALKETEI-----KEITVS-----FLOKITDLOQLKQO 106
DB 1270 EEEQRLINDLTQRLQTESGPFQRQDEKALVSQLSRGKQAFQTOIEELKQLESE 1329
QY 107 EED-----FRKQLEDEGRKAEKENTTAELTEELINKWRLLY----- 142
DB 1330 VKAKNALAHALQSSRHDCDLLREQYEEBESKAELOKALSANTVAQWRTKYETDAIOR 1389
QY 143 -EELYNKTKPFQIQLDA-----FEVEKQALLNE-----HGAAQEOQL 177

RESULT 27
Q922D2 PRELIMINARY; PRT; 1937 AA.
AC Q922D2;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Myosin heavy chain 2a.
GN MYHC-2A.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Chikuni K., Nakajima I., Muroya S.;
RL "Sequencing of the horse myosin heavy chain isoforms."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB088365; BAC05679.1; -.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin N.
DR InterPro; IPR002928; Myosin tail.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRINTS; PR00194; TROPOMYOSIN.
DR ProDom; PD000355; myosin_head; 1.
DR PROSITE; P850096; IQ; 1.
DR PROSITE; P850096; IQ; 1.
SQ SEQUENCE 1937 AA; 227247 MW; 4B00FA4246B41271 CRC64;

Query Match 15.1%; Score 180.5; DB 6; Length 1937;
Best Local Similarity 23.1%; Pred. No. 0.12;
Matches 83; Conservative 49; Mismatches 99; Indels 129; Gaps 13;

QY 8 VOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAE---DVQHQILATES 63
DB 1209 IDNLQVQKLEKEKSEMKMEIDDL-----ASNVTVSQKAGNLEKVCRTLEDQVSELKSK 1264
QY 64 NOEYVRLMLDLQTSALKETEI-----KEITVS-----FLOKITDLOQLKQO 106
DB 1265 EEEQRLINDLTQRLQTESGPFQRQDEKALVSQLSRGKQAFQTOIEELKQLESE 1324
QY 107 -----EED-----FRKQLEDEGRKAEKENTTAELTEELINKWRLLY----- 142
DB 1325 IKAKNALAHALQSSRHDCDLLREQYEEBESKAELOKALSANTVAQWRTKYETDAIOR 1384
QY 143 -EELYNKTKPFQIQLDA-----FEVEKQALLNE-----HGAAQEOQL 177
DB 1385 TEELBEAKKKLAQRLQAAEHEVAVNAKASLEKTKORLQNEVEDLMLDVERTNAACAAL 1444
QY 178 NKIRDSYAKLLG-----HQNL-----KQKIKHVVKLKDEN 207
DB 1445 DKKQNFNFKILAEWKQYETHAELEAAQKARSIGTGLFKMKNAYEESLDQLETLKRE 1504
QY 208 SOLKSEVSKLRCOLAK-----KKQ-----SETKLOEELNKVLGK 242
DB 1505 KNLQOEISDLTEQIAGGKRIHELEKIKKQVEQKSEIQAALEAEASLEHEEGKILRIQ 1564

```

```

DB 1390 TEELEBAKKLAQRLQAAEHEVAVNAKASLEKTKORLQNEVEDLMLDVERTNAACAAL 1449
QY 178 NKIRDSYAKLLG-----HQNL-----KQKIKHVVKLKDEN 207
DB 1450 DKKQNFNFKILAEWKQYETHAELEAAQKARSIGTGLFKMKNAYEESLDQLETLKRE 1509
QY 208 SOLKSEVSKLRCOLAK-----KKQ-----SETKLOEELNKVLGK 242
DB 1510 KNLQOEISDLTEQIAGGKRIHELEKIKKQVEQKSEIQAALEAEASLEHEEGKILRIQ 1569

RESULT 27
Q922D2 PRELIMINARY; PRT; 1937 AA.
AC Q922D2;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Myosin heavy chain 2a.
GN MYHC-2A.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Chikuni K., Nakajima I., Muroya S.;
RL "Sequencing of the horse myosin heavy chain isoforms."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB088365; BAC05679.1; -.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin N.
DR InterPro; IPR002928; Myosin tail.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRINTS; PR00194; TROPOMYOSIN.
DR ProDom; PD000355; myosin_head; 1.
DR PROSITE; P850096; IQ; 1.
DR PROSITE; P850096; IQ; 1.
SQ SEQUENCE 1937 AA; 227247 MW; 4B00FA4246B41271 CRC64;

Query Match 15.1%; Score 180.5; DB 6; Length 1937;
Best Local Similarity 23.1%; Pred. No. 0.12;
Matches 83; Conservative 49; Mismatches 99; Indels 129; Gaps 13;

QY 8 VOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAE---DVQHQILATES 63
DB 1209 IDNLQVQKLEKEKSEMKMEIDDL-----ASNVTVSQKAGNLEKVCRTLEDQVSELKSK 1264
QY 64 NOEYVRLMLDLQTSALKETEI-----KEITVS-----FLOKITDLOQLKQO 106
DB 1265 EEEQRLINDLTQRLQTESGPFQRQDEKALVSQLSRGKQAFQTOIEELKQLESE 1324
QY 107 -----EED-----FRKQLEDEGRKAEKENTTAELTEELINKWRLLY----- 142
DB 1325 IKAKNALAHALQSSRHDCDLLREQYEEBESKAELOKALSANTVAQWRTKYETDAIOR 1384
QY 143 -EELYNKTKPFQIQLDA-----FEVEKQALLNE-----HGAAQEOQL 177
DB 1385 TEELBEAKKKLAQRLQAAEHEVAVNAKASLEKTKORLQNEVEDLMLDVERTNAACAAL 1444
QY 178 NKIRDSYAKLLG-----HQNL-----KQKIKHVVKLKDEN 207
DB 1445 DKKQNFNFKILAEWKQYETHAELEAAQKARSIGTGLFKMKNAYEESLDQLETLKRE 1504
QY 208 SOLKSEVSKLRCOLAK-----KKQ-----SETKLOEELNKVLGK 242
DB 1505 KNLQOEISDLTEQIAGGKRIHELEKIKKQVEQKSEIQAALEAEASLEHEEGKILRIQ 1564

```

[illegible]

```

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20391973; PubMed=10931863;
RA Rogers G.C., Chui K.K., Lee E.W., Wedaman K.P., Sharp D.J.,
  Holland G., Morris R.L., Scholley J.M.;
RT "A Kinesin-related Protein, KRP(180), Positions Prometaphase Spindle
  Poles during Early Sea Urchin Embryonic Cell Division.";
RL J. Cell Biol. 150:499-512(2000).
DR EMBL: AF284333; AAG01844.1; --
DR HSP; P17119; 3KX.
DR InterPro: IPR001752; kinesin_motor.
DR Pfam: PF00225; kinesin; 1.
DR PRINTS: PR00380; KINESINHEAVY.
DR SMART: SM00129; KISC; 1.
DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE: PS00067; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 1463 AA; 166589 MW; 1E8BFC2B5AEF94D CRC64;

Query Match      15.0%; Score 179; DB 5; Length 1463;
Best Local Similarity 24.4%; Pred. No. 0.11;
Matches 59; Conservative 51; Mismatches 108; Indels 24; Gaps 6;

QY 4 YDSNVQSLSDVTAQESYKALTASIEDLKLNSLSQEKAAKAGKNAEDVQHILATSS 63
Db ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
1208 YDNRDQNE---EIRSLK-MKADELDVRISKEILQAHTALTVEIEQVRNEMAEKSS 1263

QY 64 NOEYVRML-LDLQTSALKETEL--KEITVSPFQKITDLOLQKQBEDFRKQLEDEGR 120
Db ::||| : : : : : ||| : : : ||| : : : ||| : : : ||| : : :
1264 LRDEVNHLKRDNEROKTVIASMLRDKDAVEKLYTVQTTLDQVKANBEILOENMD----- 1318

QY 121 KAEKENTTAELTEEINKWRLLYEELYNKTPFQIQIDAFVEVEKQALLNEHGAAQFOLNKI 180
Db ::||| : : : : : ||| : : : ||| : : : ||| : : : ||| : : :
1319 -----QVNEELDRTSALESTHFKEKEDIKSLEREKSKLTVDLTLKEVYEEA 1369

QY 181 RDSYAKLLGHQNLKOKIKHVVKLKDENSOLKSEVSKLRQQLAKKKQSETKQ---EELNK 237
Db :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1370 EKKITELGGHQNPQKIHHLQAVKSENYFLKEVESLEKQLGKAQSDSEQMKRDYEALQK 1429

QY 238 VL 239
Db 1430 RL 1431

```

Search completed: December 16, 2003, 06:09:25
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:07:28 ; Search time 41 Seconds

(without alignments)
936.874 Million cell updates/sec

Title: US-09-978-309A-74

Perfect score: 1195

Sequence: 1 QEKVDSMVQSLIEDVTAQFES.....KKQSETKLQBELNKVLGIK 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : A Geneseq 19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1193	99.8	646	23	Human novel polype
2	1193	99.8	709	23	Human novel polype
3	1193	99.8	725	18	Human RHAMM protei
4	1193	99.8	725	23	Human receptor for
5	1193	99.8	725	23	Human hyaluronon a
6	1178	98.6	709	23	Human novel polype
7	1105	92.5	351	17	Human umbilical ve
8	917	76.7	630	18	Mouse RHAMM protei
9	914	76.5	606	17	Receptor for hyalu

10	914	76.5	631	17	AAR999675	RHAMM 1-2a isoform
11	913	76.4	631	23	ABG50843	Mouse receptor for
12	913	76.4	794	23	AAU11437	Mouse hyaluronon a
13	893	74.7	713	23	AAU11438	Rat hyaluronon aci
14	891.5	74.6	476	14	AAR43563	Fibronectin recepto
15	210	17.6	42	22	AA91998	Human protein sequ
16	209	17.5	436	22	AA95451	Human protein sequ
17	209	17.5	1388	23	AAU79590	Human kinesin moto
18	209	17.5	1388	23	AAE14400	Human kinesin supe
19	209	17.5	1388	24	ABR48222	Human bladder can
20	191	16.0	795	23	ABR77430	Human tumour marke
21	191	16.0	976	22	AAG66581	Human SCP-1 mutein
22	191	16.0	976	24	ABP74709	Human SCP-1 protei
23	191	16.0	1411	17	AAW02258	Nucleolar/endosoma
24	189	15.8	1690	22	ABB61144	Drosophila melanog
25	189	15.8	1690	22	ABB61173	Drosophila melanog
26	179.5	15.0	1948	22	ABG21233	Novel human diagno
27	178.5	14.9	1940	23	ABG79661	Invertebrate forag
28	177	14.8	717	21	AA21231	Tomato LEMPP1. Ly
29	171.5	14.4	336	23	ABT78804	Myosin tail Myosn
30	170	14.2	1388	23	ABB57354	Mouse ischaemic co
31	169	14.1	951	22	ABUS3070	Intracellular traf
32	169	14.1	953	22	ABUS3069	Intracellular traf
33	169	14.1	953	22	ABUS3071	Intracellular traf
34	169	14.1	961	22	ABUS3077	Intracellular traf
35	169	14.1	1881	23	ABP73809	Candida albicans e
36	169	14.1	2230	24	ABU07445	Protein differenti
37	164	13.7	551	22	ABAB3244	Human protein sequ
38	164	13.7	1213	22	AA40016	Human polypeptide
39	164	13.7	1935	23	ABG31649	Amino acid distrib
40	163.5	13.7	1017	22	AAE02246	Domestic mite Bll1
41	163	13.6	561	19	AAW63043	Streptococcus uber
42	163	13.6	1318	23	ABB77985	Amino acid sequenc
43	162.5	13.6	853	23	ABAB3794	Herbicidally activ
44	162.5	13.6	1939	23	ABB77096	Human alpha-myosin
45	162	13.6	1489	22	ABBS9948	Drosophila melanog
46	161.5	13.5	1886	19	AAW54241	Rattus norvegicus
47	160.5	13.4	1286	21	AAAB43359	Human ORFX ORF3123
48	160.5	13.4	1286	24	AAO26961	Human CRK related
49	160.5	13.4	1958	23	ABB81928	Human kinase #2.
50	160.5	13.4	2053	22	AAU03501	Human protein kina

ALIGNMENTS

RESULT 1
ABG70292
ID ABG70292 standard; Protein; 646 AA.
XX
AC ABG70292;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human novel polypeptide #8.
XX
KW Human; trauma; viral infection; parasitic infection; addiction;
KW bacterial infection; Alzheimer's disease; Huntington's disease; anxiety;
KW Parkinson's disease; behavioural disorder; pain; hair growth disease;
KW alopecia; pigmentation disorder; inflammatory disorder; arthritis; AIDS;
KW inflammatory bowel disease; Crohn's disease; cancer; adenocarcinoma;
KW acquired immunodeficiency syndrome; colon; asthma; hypertension; obesity;
KW autoimmune disease; diabetes; graft versus host disease; ulcer; bulimia;
KW anorexia; dementia; gene therapy.

17-DEC-2001; 2001WO-US49122.

```

PR 15-DEC-2000; 2000US-256025P.
PR 30-JAN-2001; 2001US-265163P.
PR 02-MAR-2001; 2001US-272929P.
PR 09-MAR-2001; 2001US-274864P.
PR 16-MAR-2001; 2001US-276688P.
PR 22-MAR-2001; 2001US-277880P.
PR 25-APR-2001; 2001US-286409P.
PR 31-JUL-2001; 2001US-309246P.
PR 29-AUG-2001; 2001US-315600P.
XX (CURA-) CURAGEN CORP.
XX Shinkets RA, Colman SD, Spytek KA, Ballinger RA, Guo X;
XX Tchernev VI, Shenoy SG, Li L, Ellerman KE, Zerhusen BD;
PI Patturajan M, Casman SJ, Boldog F, Gusev VY, Burgess CE;
PI Edinger S, Gangolli EA, Malyankar UM, Gunther E, Smithson G;
PI Millet I, Gerlach VL;
XX WPI; 2002-590743/63.
DR N-PSDB; ABS51735.
XX Novel polypeptide, designated NOVX for treating or preventing disorders
PT or symptoms e.g. trauma, Alzheimer's disease, cancers, acquired
PT immunodeficiency syndrome, asthma and rheumatoid arthritis -
XX Claim 1; Page 46; 252pp; English.
XX The invention relates to human novel polynucleotides and polypeptides.
XX The sequences are useful for the treatment, prevention and diagnosis of
CC disorders such as trauma, viral/parasitic/bacterial infections,
CC Alzheimer's disease, Huntington's disease, Parkinson's disease,
CC behavioural disorders, anxiety, addiction, pain, hair growth diseases,
CC alopecia, pigmentation disorder, inflammatory disorders such as osteo-
CC and rheumatoid arthritis, inflammatory bowel disease, Crohn's disease,
CC acquired immunodeficiency syndrome (AIDS), cancers such as colon cancer
CC and adenocarcinoma, asthma, hypertension, autoimmune disease, diabetes,
CC obesity, graft versus host disease, ulcer, bulimia, anorexia and
CC dementia. Sequences ABG70285-ABG70305 represent human novel polypeptides
CC of the invention.
XX WPI; 2002-590743/63.
DR N-PSDB; ABS51735.
XX Sequence 646 AA;
SQ Query Match 99.8%; Score 1193; DB 23; Length 646;
PT Best Local Similarity 99.8%; Pred. No. 3.5e-78;
PT Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 QEKYDSMVQSLDVTQAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 60
DB 364 QEKYDSMVQSLDVTQAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 423
QY 61 ESSNQBYVRMLLDLQTKSALKETEIKETVTSFLQKITDQNLQKQOEDFRKQLEDEGR 120
DB 424 ESSNQBYVRMLLDLQTKSALKETEIKETVTSFLQKITDQNLQKQOEDFRKQLEDEGR 483
QY 121 KAEKNTTAEITBEINKWRLLYEELYNKTKPFQIQLDAPEVEKQALLNHEGAAQOQLNKI 180
DB 484 KAEKNTTAEITBEINKWRLLYEELYNKTKPFQIQLDAPEVEKQALLNHEGAAQOQLNKI 543
QY 181 RDSYAKLGHQNLKQIKHVVKLKDENSQKLSVSKLRCLQAKKQSETKLQBELNKVLG 240
DB 544 RDSYAKLGHQNLKQIKHVVKLKDENSQKLSVSKLRCLQAKKQSETKLQBELNKVLG 603
QY 241 IK 242
DB 604 IK 605
XX
RESULT 2
ABG70291
ID ABG70291 standard; Protein; 709 AA.
XX
AC ABG70291;
XX

```

```

DT 21-OCT-2002 (first entry)
XX Human novel polypeptide #7.
DE XX Human; trauma; viral infection; parasitic infection; addiction;
XX Human; bacterial infection; Alzheimer's disease; Huntington's disease; anxiety;
XX Parkinson's disease; behavioural disorder; pain; hair growth disease;
XX alopecia; pigmentation disorder; inflammatory disorder; arthritis; AIDS;
XX inflammatory bowel disease; Crohn's disease; cancer; adenocarcinoma;
XX acquired immunodeficiency syndrome; colon; asthma; hypertension; obesity;
XX autoimmune disease; diabetes; graft versus host disease; ulcer; bulimia;
XX anorexia; dementia; gene therapy.
XX Homo sapiens.
OS WO200257452-A2.
XX 25-JUL-2002.
PD 17-DEC-2001; 2001WO-US49122.
XX 15-DEC-2000; 2000US-256025P.
XX 30-JAN-2001; 2001US-265163P.
PR 02-MAR-2001; 2001US-272929P.
PR 09-MAR-2001; 2001US-274864P.
PR 16-MAR-2001; 2001US-276688P.
PR 22-MAR-2001; 2001US-277880P.
PR 25-APR-2001; 2001US-286409P.
PR 31-JUL-2001; 2001US-309246P.
PR 29-AUG-2001; 2001US-315600P.
XX (CURA-) CURAGEN CORP.
PA Shinkets RA, Colman SD, Spytek KA, Ballinger RA, Guo X;
XX Tchernev VI, Shenoy SG, Li L, Ellerman KE, Zerhusen BD;
PI Patturajan M, Casman SJ, Boldog F, Gusev VY, Burgess CE;
PI Edinger S, Gangolli EA, Malyankar UM, Gunther E, Smithson G;
PI Millet I, Gerlach VL;
XX WPI; 2002-590743/63.
DR N-PSDB; ABS51734.
XX Novel polypeptide, designated NOVX for treating or preventing disorders
PT or symptoms e.g. trauma, Alzheimer's disease, cancers, acquired
PT immunodeficiency syndrome, asthma and rheumatoid arthritis -
XX Claim 1; Page 45; 252pp; English.
XX The invention relates to human novel polynucleotides and polypeptides.
XX The sequences are useful for the treatment, prevention and diagnosis of
CC disorders such as trauma, viral/parasitic/bacterial infections,
CC Alzheimer's disease, Huntington's disease, Parkinson's disease,
CC behavioural disorders, anxiety, addiction, pain, hair growth diseases,
CC alopecia, pigmentation disorder, inflammatory disorders such as osteo-
CC and rheumatoid arthritis, inflammatory bowel disease, Crohn's disease,
CC acquired immunodeficiency syndrome (AIDS), cancers such as colon cancer
CC and adenocarcinoma, asthma, hypertension, autoimmune disease, diabetes,
CC obesity, graft versus host disease, ulcer, bulimia, anorexia and
CC dementia. Sequences ABG70285-ABG70305 represent human novel polypeptides
CC of the invention.
XX WPI; 2002-590743/63.
DR N-PSDB; ABS51734.
XX Sequence 709 AA;
SQ Query Match 99.8%; Score 1193; DB 23; Length 709;
PT Best Local Similarity 99.6%; Pred. No. 3.9e-78;
PT Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 QEKYDSMVQSLDVTQAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 60
DB 427 QEKYDSMVQSLDVTQAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 486
QY 61 ESSNQBYVRMLLDLQTKSALKETEIKETVTSFLQKITDQNLQKQOEDFRKQLEDEGR 120

```

Dd 487 ESSNQYVRLMLDLQTKSALKETEIKETVTSFLQKITDQLNQKQOEEDFRKQLEDEGR 546
 QY 121 KAEKNTTAEITTEINKWLLYELYNKTKPFIQIDAFEVEKQALLNHHGAQOEOLNKI 180
 Dd 547 KAEKNTTAEITTEINKWLLYELYNKTKPFIQIDAFEVEKQALLNHHGAQOEOLNKI 606
 QY 181 RDSYAKLLGHQNLKOKIKHVVKLKDENSOLKSEVSKLRCLAKKQSETKLOEELNKVLG 240
 Dd 607 RDSYAKLLGHQNLKOKIKHVVKLKDENSOLKSEVSKLRCLAKKQSETKLOEELNKVLG 666
 QY 241 IK 242
 Dd 667 IK 668
 RESULT 3
 AAW39165
 ID AAW39165 standard; Protein; 725 AA.
 XX AC AAW39165;
 XX DT 27-APR-1998 (first entry)
 XX DE Human RHAMM protein.
 XX KW Hyaluronan receptor; receptor for hyaluronic acid mediated motility;
 KW RHAMM; glycosaminoglycan; binding domain; human; oncogene; treatment;
 KW growth factor; cell locomotion disorder; dementia; detection;
 KW inflammatory disorder; autoimmune disease; diagnosis; prognosis.
 XX OS Homo sapiens.
 XX FH Key
 FH Region 478..480
 FT /note="Region indicated in specification"
 FT Region 588..570
 FT /note="Region indicated in specification"
 FT Region 589..591
 FT /note="Region indicated in specification"
 FT Binding-site 636..646
 FT /note="Hyaluronic acid binding domain"
 FT Binding-site 658..667
 FT /note="Hyaluronic binding domain"
 XX PN WO9738098-A1.
 XX PD 16-OCT-1997.
 XX PF 10-APR-1997; 97WO-CA00240.
 XX PR 10-APR-1996; 96GB-0007441.
 XX PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
 XX PA (UYMA-) UNIV MANITOBA.
 XX PI Entwistle J, Turley EA;
 XX WPI; 1997-512715/47.
 DR N-PSDB; AAV02800.
 XX PT Isolated human receptor for hyaluronic acid mediated motility - used
 PT to develop products for treating e.g. tumours, inflammatory
 PT disorders, dementia, AIDS, diabetes and auto-immune diseases
 XX PS Claim 16; Page 46; 66pp; English.
 XX CC This sequence represents the human hyaluronan receptor which is also
 CC known as the receptor for hyaluronic acid mediated motility (RHAMM).
 CC Hyaluronan is a large glycosaminoglycan that is ubiquitous in the
 CC extracellular matrix and whose synthesis has been linked to cell
 CC migration, growth and transformation. It interacts with cell surfaces via
 CC specific protein receptors, e.g. RHAMM, that mediate many biological
 CC effects. The RHAMM/Hyaluronic acid interaction is involved in

CC oncogene-and growth factor-mediated cell locomotion. The products can be
 CC used in the treatment of disorders involving cell locomotion, e.g. tumour
 CC invasion, birth defects, acute and chronic inflammatory disorders,
 CC Alzheimer's and other forms of dementia, including Parkinson's and
 CC Huntington's diseases, AIDS, diabetes, autoimmune diseases, corneal
 CC dysplasia and hypertrophies, burns, surgical incisions and adhesions,
 CC strokes and multiple sclerosis. They can also be used in e.g. CNS and
 CC spinal cord regeneration, contraception and in vitro fertilisation and
 CC embryo development. The products can also be used in detection, diagnosis
 CC and prognosis.
 XX SQ Sequence 725 AA;
 Query Match 99.8%; Score 1193; DB 18; Length 725;
 Best Local Similarity 99.6%; Pred. No. 4e-78;
 Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CEKYDSMVQSLDVTAFQESYKALTASIEDLKLNSSLOEKAAGKNAEDVQHQLLAT 60
 Dd 443 CEKYDSMVQSLDVTAFQESYKALTASIEDLKLNSSLOEKAAGKNAEDVQHQLLAT 502
 QY 61 ESSNQYVRLMLDLQTKSALKETEIKETVTSFLQKITDQLNQKQOEEDFRKQLEDEGR 120
 Dd 503 ESSNQYVRLMLDLQTKSALKETEIKETVTSFLQKITDQLNQKQOEEDFRKQLEDEGR 562
 QY 121 KAEKNTTAEITTEINKWLLYELYNKTKPFIQIDAFEVEKQALLNHHGAQOEOLNKI 180
 Dd 563 KAEKNTTAEITTEINKWLLYELYNKTKPFIQIDAFEVEKQALLNHHGAQOEOLNKI 622
 QY 181 RDSYAKLLGHQNLKOKIKHVVKLKDENSOLKSEVSKLRCLAKKQSETKLOEELNKVLG 240
 Dd 623 RDSYAKLLGHQNLKOKIKHVVKLKDENSOLKSEVSKLRCLAKKQSETKLOEELNKVLG 682
 QY 241 IK 242
 Dd 683 IK 684
 RESULT 4
 ABG60842
 ID ABG60842 standard; Protein; 725 AA.
 XX AC ABG60842;
 XX DT 13-AUG-2002 (first entry)
 XX DE Human receptor for hyaluronan-mediated motility (RHAMM).
 XX KW Tissue disorder; response-to-injury process; cell proliferating;
 KW hyaluronic acid; HA; receptor for hyaluronan-mediated motility;
 KW RHAMM; inflammatory neurological disorder; Parkinson's disease;
 KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;
 KW inflammatory dermatosis; psoriasis; inflammatory bowel disease;
 KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;
 KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;
 KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;
 KW tissue transplantation; stroke; inflammatory response; fibrotic response;
 KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;
 KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;
 KW septic shock; thyroiditis; retinopathy.
 XX OS Homo sapiens.
 XX PN WO200228415-A1.
 XX PD 11-APR-2002.
 XX PF 05-OCT-2000; 2000WO-IB01534.
 XX PR 05-OCT-2000; 2000WO-IB01534.
 XX PA (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.

```
PI Tuxley EA, Cruz TF;
XX WPI; 2002-435298/46.
XX
XX Treating tissue disorder associated with response-to-injury process or
PT proliferating cells in mammals, e.g. fibrosis, inflammation, by
PT administering a compound that alters activity of transition molecules
PT within a cell
XX
XX Example 30; Fig 50; 215pp; English.
XX
XX The invention describes a method of treating a tissue disorder associated
CC with response-to-injury process or proliferating cells in a patient,
CC comprising administering a polypeptide (I) which binds hyaluronic acid
CC (HA), an antibody which binds one of domains D1-D5 of Receptor for
CC hyaluronan-mediated motility (RHAMM), a polypeptide fragment encoding
CC any of D1-D5 of RHAMM, or a vector which expresses antisense RHAMM,
CC antibodies or a polypeptide fragment. The method is useful for treating a
CC patient with an inflammatory neurological disorder such as Parkinson's
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),
CC inflammatory bowel disease, stenosis or restenosis, cancer, kidney
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic
CC fibrosis), obesity or obesity related diseases, lupus, cardiovascular
CC disease (e.g. atherosclerosis), and wound especially surgical excision
CC adhesions, to prevent scar and also for treating or preventing diabetes
CC mellitus. The method is also useful for treating tissue transplantation
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response
CC associated with medical implants such as hip implants, vascular wraps and
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence
CC represents a receptor for hyaluronan-mediated motility protein used in
CC the method of treating a tissue disorder described in the invention.
XX
XX Sequence 725 AA;
SQ
Query Match 99.8%; Score 1193; DB 23; Length 725;
Best Local Similarity 99.6%; Pred. No. 4e-78; Indels 0; Gaps 0;
Matches 241; Conservative 1; Mismatches 0;
QY 1 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 60
Db 443 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 502
QY 61 ESSNQEVYRMLLDLQTKSALKETEIKETVTSFLQKITDQLQNLKQOEEDPRKQLEDEGR 120
Db 503 ESSNQEVYRMLLDLQTKSALKETEIKETVTSFLQKITDQLQNLKQOEEDPRKQLEDEGR 562
QY 121 KAKENTTAELTEINKWRLLYBELYNKTKPFQILDAPFEVQKALLNEHGAQEQLNKI 180
Db 563 KAKENTTAELTEINKWRLLYBELYNKTKPFQILDAPFEVQKALLNEHGAQEQLNKI 622
QY 181 RDSYAKLLGHONLKQIKHVVKLKDENSOLKSEVSKLRCOLAKKQSETKLQELKNKVLG 240
Db 623 RDSYAKLLGHONLKQIKHVVKLKDENSOLKSEVSKLRCOLAKKQSETKLQELKNKVLG 682
QY 241 IK 242
Db 683 IK 684
RESULT 5
AAU11436
ID AAU11436 standard; Protein; 725 AA.
XX
XX AAU11436;
AC
XX
XX 12-MAR-2002 (first entry)
DT
XX
XX Human hyaluronic acid binding protein RHAMM.
DE
XX
XX Human; hyaluronic acid binding protein; RHAMM; gene therapy;
KW receptor for HA mediated mobility; immunosuppressive; cytostatic.
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Misc-difference 278 /note= "Encoded by GAA"
FT Misc-difference 299 /note= "Encoded by AAA"
FT Misc-difference 323 /note= "Encoded by AAA"
FT Misc-difference 331 /note= "Encoded by CAG"
XX
XX WO200180899-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-CA00533.
XX 20-APR-2000; 2000US-198613P.
XX (CANG-) CANGENE CORP.
XX
XX Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;
XX WPI; 2002-075094/10.
XX N-PSDB; AAS17496.
XX
XX Protein conjugates that selectively target certain tissues and organs
XX useful for treating and preventing various diseases, comprises
XX glucose-aminoglycan-targeting domain conjugated to a therapeutic
XX protein
XX
XX Claim 6; Page 115; 121pp; English.
XX
XX The invention relates to a conjugate comprising an hyaluronic acid (HA)
XX -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
XX contiguous with, or coupled to a polypeptide conjugated to a therapeutic
XX agent, and the polynucleotides encoding them. Also included is a method
XX for preparation of the HA-binding protein by inserting a first nucleotide
XX sequence encoding a HA-binding protein directly linked to a second
XX nucleotide sequence encoding a therapeutic protein into a suitable
XX vector, expressing the vector in an acceptable host, purifying conjugate
XX molecule from host or expression medium. The composition is useful for
XX altering in vivo the distribution of a therapeutic agent comprising
XX administering the composition to the animal where conjugate molecule will
XX distribute primarily in tissues and organs containing high levels of
XX endogenous HA and for treating mammal with a disorder where a diseased
XX tissue of the mammal contains high level of HA e.g. rheumatoid
XX arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
XX dosages required also translates into lower immunogenicity of the
XX conjugated protein as compared to the native protein. As a result,
XX conjugates improve patient compliance and reduce direct and indirect
XX costs associated with the drug substance and its administration.
XX Conjugates allows for the use, where appropriate, of lower, safer,
XX dosages as compared to the conventional dosage requirements for the
XX unconjugated corresponding therapeutic agent. Conjugate molecules have an
XX increased half-life and potency, resulting in prolonged circulation of
XX the molecule, efficient distribution into the target tissues, and
XX increased bioavailability. The present sequence represents a RHAMM
XX protein.
XX
XX Sequence 725 AA;
Query Match 99.8%; Score 1193; DB 23; Length 725;
Best Local Similarity 99.6%; Pred. No. 4e-78; Indels 0; Gaps 0;
Matches 241; Conservative 1; Mismatches 0;
QY 1 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 60
Db 443 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHILAT 502
```



```

FT Modified-site 91...93
FT /label= N-glycosylation_site
FT 258..260
FT Modified-site
FT /label= N-glycosylation_site
FT 279..281
FT Modified-site
FT /label= N-glycosylation_site
FT 300..302
FT Modified-site
FT /label= N-glycosylation_site
FT 321..323
FT Modified-site
FT /label= N-glycosylation_site
FT 342..344
FT Modified-site
FT /label= N-glycosylation_site
FT 373..375
FT Modified-site
FT /label= N-glycosylation_site
FT 413..415
FT Modified-site
FT /label= N-glycosylation_site
FT 532..542
FT Domain
FT /label= Hyaluronan_binding_domain-I
FT 553..562
FT Domain
FT /label= Hyaluronan_binding_domain-I
FT 594..596
FT Modified-site
FT /label= N-glycosylation_site
FT
FT EP721012-A2
FT
FT 10-JUL-1996
FT
FT 16-OCT-1995; 95EP-0307310
FT
FT 14-OCT-1994; 94GB-0020740
FT
FT (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
FT (UYMA-) UNIV MANITOBA.
FT
FT Entwistle J, Turley EA, Zhang S;
FT
FT WPI; 1996-310997/32.
FT N-PSDB; AAT34499, AAT34500, AAT23423, AAT34524.
FT
FT Receptor for hyaluronic acid-mediated motility protein, and DNA
FT encoding it - useful to treat or prevent diseases associated with
FT the receptor, e.g. breast cancer
FT
FT Claim 2; Page 40-42; 117pp; English.
FT
FT RHAMM 1 (AAR99673), or Receptor for Hyaluronic Acid Mediated Motility,
FT is a hyaluronan receptor protein which is involved in cell locomotion
FT or motility and cell proliferation and transformation. Its amino
FT acid sequence was deduced from a cDNA clone (AAT34499) obt'd. from
FT murine 3T3 cells and a genomic sequence (AAT34500) from a mouse
FT fibroblast genomic library. 2 Alternative mRNAs for RHAMM 1 were
FT found, RHAMM 1A (see also AAT34523) and RHAMM 1B (AAT34524), that had
FT identical translated portions. Increased expression of RHAMM 1
FT protein is indicative of a poor prognosis for breast cancer. The
FT protein can be used to suppress or control a tumour by modulating
FT the interaction of cell-associated RHAMM with its ligand.
FT
FT SQ Sequence 606 AA;
FT
FT Query Match 76.5%; Score 914; DB 17; Length 606;
FT Best Local Similarity 76.0%; Pred. No. 4.9e-58;
FT Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;
FT
FT QY 1 QEKYDSMVQSLEDVTAQFESYKALTASEIEDLKLENSLOEKAAGNAEDVQHILAT 60
FT
FT Db
FT 338 QEKYNDTAQSLRDVTAQLESYKSLFKETEDLKLENTLQEKVMAEKSVEDVQQIILTA 397
FT
FT QY 61 ESSNQEVYRMLLDQTKALKETEKEITVSLQKITDLQNLQKQEDDFRQLEDEGR 120
FT
FT Db 398 ESTNQEVARMVDLQNRSLTKEEIEKETSFLKXITDLQNLQROQDDFRQLEEKGR 457
FT
FT QY 121 KAEKENTTAETEEINKVRLLYEELYNKTKPQIQLDAFEVEKQALLNEHGAQOELNKI 180

```

```

Db 458 TAEKENVMTETLWEINKWRLLYDELYETKTFPQQQLDAFEAEKQALLNEHGAQOELNKI 517
QY 181 RDSYAKLGHQNLKQKIKHVVKLKDENSQKSEVSKLRCOLAKKQKQSETKLOBELNKVLG 240
Db 518 RDSYAQLGHQNLKQKIKHVVKLKDENSQKSEVSKLRSQKVRKQNELRQGBLDKALG 577
QY 241 IK 242
Db 578 IR 579

RESULT 10
AAR99675
ID AAR99675 standard; Protein; 631 AA.
XX
AC AAR99675;
XX
DT 10-OCT-1996 (first entry)
XX
DE RHAMM 1-2a isoform.
XX
KW RHAMM 1-2a; receptor for hyaluronic acid mediated motility;
KW hyaluronan receptor; cell locomotion; cell proliferation;
KW breast cancer; therapy.
XX
OS Mus sp.
XX
Key Location/Qualifiers
FT Region 55..79
FT FT /note= "exon 2A-encoded region"
XX
FN EP721012-A2.
XX
PD 10-JUL-1996.
XX
PF 16-OCT-1995; 95EP-0307310.
XX
PR 14-OCT-1994; 94GB-0020740.
XX
PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
PA (UYMA-) UNIV MANITOBA.
XX
PI Entwistle J, Turley EA, Zhang S;
XX
DR WPI; 1996-310997/32.
DR N-PSDB; AAT34525.
XX
PT Receptor for hyaluronic acid-mediated motility protein, and DNA
PT encoding it - useful to treat or prevent diseases associated with
PT the receptor, e.g. breast cancer
XX
PS Claim 8; Page 50-52; 117pp; English.
XX
RHAMM 1-2a (AAR99675) is an alternatively spliced variant of
RHAMM 1 (AAR99673) (receptor for hyaluronic acid mediated
motility), a protein involved in cell locomotion or motility and
cell proliferation and transformation. It differs from RHAMM 1
by an insertion of 25 amino acids (see also AAR99674) between
amino acids 54 and 55 of RHAMM 1, resulting from an alternative
spliced exon 2A (AAT34502). RHAMM 1-2a is the isoform that is
overexpressed in tumours. Determination of the level of RHAMM
1-2a in a sample can be used to assess the prognosis of a tumour
(esp. breast cancer) patient. The RHAMM 1-2a protein can also
be used to suppress or control a tumour by modulating the
interaction of cell-associated RHAMM with its ligand.
XX
SQ Sequence 631 AA;
XX
Query Match 76.5%; Score 914; DB 17; Length 631;
Best Local Similarity 76.0%; Pred. No. 5.1e-58;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;
QY 1 QEKYDSMVQSLEDVTAQFESYKALTASEIEDLKLENSLOEKAAGNAEDVQHILAT 60

```

Db 363 QEKYNDTAQSRDVTVAQLESYKSTLKEIEDLKLENLTLOEKVAAEKSVEDVQQILTA 422
Qy 61 ESSNQEVYRMLLDLQTSKALKEITEITVSFLOKITDLOLQKQOEEDFRKQLEDEGR 120
Db 423 ESTNQEVYARMVQDLQNRSTLKEEIEKETSFLKIDTLKNLQKQOEEDFRKQLEEKGR 482
Qy 121 KAEKENTTAELTEINKRWLLYEBLYNKTPFQOLDFAFEVKEQALLNEHGAAGQOLNKI 180
Db 483 TAEKENVMTLMEINKRWLLYEBLYNKTPFQOLDFAFEVKEQALLNEHGAAGQOLNKI 542
Qy 181 RDSYAKLLGHQNLQKIKHVVKLKDENSQKSEVSKLRQCLAKKQKQSTKLOEBLNKVLG 240
Db 543 RDSYAQLLGHQNLQKIKHVVKLKDENSQKSEVSKLRQCLAKKQKQSTKLOEBLNKVLG 602
Qy 241 IK 242
Db 603 IR 604

RESULT 11
ABG60843
ID ABG60843 standard; Protein; 631 AA.
XX AC ABG60843;
XX DT 13-AUG-2002 (first entry)
XX DE Mouse receptor for hyalauronan-mediated motility (RHAMM).
XX KW Tissue disorder; response-to-injury process; cell proliferating;
KW RHAMM; inflammatory neurological disorder; Parkinson's disease;
KW Alzheimer's disease; arthritis; multiple sclerosis; gastritis; nephritis;
KW inflammatory dermatosis; psoriasis; inflammatory bowel disease;
KW stenosis; restenosis; cancer; kidney fibrosis; inflammatory lung disease;
KW emphysema; asthma; cystic fibrosis; obesity; obesity related disease;
KW lupus; cardiovascular disease; atherosclerosis; wound; scar; diabetes;
KW tissue transplantation; stroke; inflammatory response; fibrotic response;
KW medical implant; Acquired immunodeficiency syndrome; AIDS; hepatitis;
KW myocardial fibrosis; hepatic fibrosis; chronic cystitis; acute mastitis;
KW septic shock; thyroiditis; retinopathy.
XX OS Mus musculus.
XX FN WO200228415-A1.
XX PD 11-APR-2002.
XX PF 05-OCT-2000; 2000WO-IB01534.
XX PR 05-OCT-2000; 2000WO-IB01534.
XX PA (TRAN-) TRANSITION THERAPEUTICS & DIAGNOSTICS IN.
XX PI Turley EA, Cruz TF;
XX DR WPI; 2002-435298/46.
XX PT Treating tissue disorder associated with response-to-injury process or
PT proliferating cells in mammals, e.g. fibrosis, inflammation, by
PT administering a compound that alters activity of transition molecules
PT within a cell -
XX PS Disclosure; Fig 50; 215pp; English.
XX CC The invention describes a method of treating a tissue disorder associated
CC with response-to-injury process or proliferating cells in a patient,
CC comprising administering a polypeptide (I) which binds hyaluronic acid
CC (HA), an antibody which binds one of domains D1-D5 of Receptor for
CC hyalauronan-mediated motility (RHAMM), a polypeptide fragment encoding
CC any of D1-D5 of RHAMM, or a vector which expresses antisense RHAMM,
CC antibodies or a polypeptide fragment. The method is useful for treating a

CC patient with an inflammatory neurological disorder such as Parkinson's
CC disease, Alzheimer's disease, arthritis including rheumatoid arthritis,
CC osteoarthritis, multiple sclerosis, inflammatory dermatosis (psoriasis),
CC inflammatory bowel disease, stenosis or restenosis, cancer, kidney
CC fibrosis, inflammatory lung disease (e.g. emphysema, asthma, cystic
CC fibrosis), obesity or obesity related diseases, lupus, cardiovascular
CC disease (e.g. atherosclerosis), and wound especially surgical excision
CC adhesions, to prevent scar and also for treating or preventing diabetes
CC mellitus. The method is also useful for treating tissue transplantation
CC (e.g. skin grafts), stroke, inflammatory responses or fibrotic response
CC associated with medical implants such as hip implants, vascular wraps and
CC catheters), inflammatory diseases such as AIDS, myocardial and hepatic
CC fibrosis, chronic cystitis, acute mastitis, gastritis, nephritis,
CC hepatitis, septic shock, thyroiditis, and retinopathy. This sequence
CC represents a receptor for hyalauronan-mediated motility protein used in
CC the method of treating a tissue disorder described in the invention.
XX SQ Sequence 631 AA;
Query Match 76.4%; Score 913; DB 23; Length 631;
Best Local Similarity 76.0%; Pred. No. 6.1e-58;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;
Qy 1 QEKYDSMVQSLIEDVTAQFESYKALTASEIEDLKLENSLQEKAKAGNAEDVQHILAT 60
Db 363 QEKYNDTAQSRDVSQALESYKSTLKEIEDLKLENLTLOEKVAAEKSVEDVQQILTA 422
Qy 61 ESSNQEVYRMLLDLQTSKALKEITEITVSFLOKITDLOLQKQOEEDFRKQLEDEGR 120
Db 423 ESTNQEVYARMVQDLQNRSTLKEEIEKETSFLKIDTLKNLQKQOEEDFRKQLEEKGR 482
Qy 121 KAEKENTTAELTEINKRWLLYEBLYNKTPFQOLDFAFEVKEQALLNEHGAAGQOLNKI 180
Db 483 TAEKENVMTLMEINKRWLLYEBLYNKTPFQOLDFAFEVKEQALLNEHGAAGQOLNKI 542
Qy 181 RDSYAKLLGHQNLQKIKHVVKLKDENSQKSEVSKLRQCLAKKQKQSTKLOEBLNKVLG 240
Db 543 RDSYAQLLGHQNLQKIKHVVKLKDENSQKSEVSKLRQCLAKKQKQSTKLOEBLNKVLG 602
Qy 241 IK 242
Db 603 IR 604
RESULT 12
AAU11437
ID AAU11437 standard; Protein; 794 AA.
XX AC AAU11437;
XX DT 12-MAR-2002 (first entry)
XX DE Mouse hyaluronic acid binding protein RHAMM.
XX KW Mouse; hyaluronic acid binding protein; RHAMM; gene therapy;
KW receptor for HA mediated motility; immunosuppressive; cytostatic.
KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
XX OS Mus musculus.
XX FH Key Location/Qualifiers
FT Misc-difference 55 /note= "Encoded by AAC"
FT Misc-difference 71 /note= "Encoded by ACG"
FT Misc-difference 89 /note= "Encoded by CAA"
FT Misc-difference 91 /note= "Encoded by CAC"
FT Misc-difference 540 /note= "Encoded by ACT"
FT Misc-difference 668 /note= "Encoded by GAT"


```
XX PN WO200180899-A2.
XX XX
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-CA00533.
XX XX
XX PR 20-APR-2000; 2000US-198613P.
XX PA (CANG-) CANGENE CORP.
XX PI Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;
XX DR WPI; 2002-075094/10.
XX DR N-PSDB; AAS17497.
XX XX
XX PT Protein conjugates that selectively target certain tissues and organs
XX PT useful for treating and preventing various diseases, comprises
XX PT glucose-aminoglycan-targeting domain conjugated to a therapeutic
XX PT protein -
XX XX
XX PS Claim 6; Page 116; 121pp; English.
XX CC The invention relates to a conjugate comprising an hyaluronic acid (HA)
XX CC -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
XX CC contiguous with, or coupled to a polypeptide conjugated to a therapeutic
XX CC agent, and the polynucleotides encoding them. Also included is a method
XX CC for preparation of the HA-binding protein by inserting a first nucleotide
XX CC sequence encoding a HA-binding protein directly linked to a second
XX CC nucleotide sequence encoding a therapeutic protein into a suitable
XX CC vector, expressing the vector in an acceptable host, purifying conjugate
XX CC molecule from host or expression medium. The composition is useful for
XX CC altering in vivo the distribution of a therapeutic agent comprising
XX CC administering the composition to the animal where conjugate molecule will
XX CC distribute primarily in tissues and organs containing high levels of
XX CC endogenous HA and for treating mammal with a disorder where a diseased
XX CC tissue of the mammal contains high level of HA e.g. rheumatoid
XX CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
XX CC dosages required also translates into lower immunogenicity of the
XX CC conjugated protein as compared to the native protein. As a result,
XX CC conjugates improve patient compliance and reduce direct and indirect
XX CC costs associated with the drug substance and its administration.
XX CC Conjugates allows for the use, where appropriate, of lower, safer,
XX CC dosages as compared to the conventional dosage requirements for the
XX CC unconjugated corresponding therapeutic agent. Conjugate molecules have an
XX CC increased half-life and potency, resulting in prolonged circulation of
XX CC the molecule, efficient distribution into the target tissues, and
XX CC increased bioavailability. The present sequence represents a RHAMM
XX CC protein.
XX SQ Sequence 794 AA;
Query Match 76.4%; Score 913; DB 23; Length 794;
Best Local Similarity 76.0%; Pred. No. 8e-58;
Matches 184; Conservative 20; Mismatches 38; Indels 0; Gaps 0;
Qy 1 QEKYDSMVQSLDVTQAFESKVALTASIEDLKLNSLSQKAAKAGNADVDVQHILAT 60
Db 526 QEKYNDTAQSLRDVSAQLESYKSTLKEIDELKLNLTLQEKVAAEKSVEDVQOQILTA 585
Qy 61 ESSNOEYVRMLDLOTGALKETEIKETITVGFLOKITDLQNLKQEDDFRQKLEDEGR 120
Db 586 ESTNOEYARMVDLQNRSLTKEEIKETITSSFLEXITDLQNLKQEDDFRQKLEEKGR 645
Qy 121 KAEKENTTAETETINKRWLYEELYNKTKPQIQIDAFVEVEKQALLNEHGAAGQNLNKI 180
Db 646 TAEKENVMTLTWEINKRWLYEELYETKTPQQQLDAFAEKQALLNEHGATQQLNKI 705
Qy 181 RDSYAKLIGHQNLKOKIKHVVKLKDENSQKSEVSKLQCLAKKQKQSETKQBELNKVLG 240
Db 706 RDSYAQLLGHQNLKOKIKHVVKLKDENSQKSEVSKLRSQVLRKQNELRLQGLDKALG 765
Qy 241 IK 242
```

```
Db 766 IR 767
RESULT 13
AAU11438
ID AAU11438 standard; Protein; 713 AA.
XX AC AAU11438;
XX DT 12-MAR-2002 (first entry)
XX DE Rat hyaluronic acid binding protein RHAMM.
XX KW Rat; hyaluronic acid binding protein; RHAMM; gene therapy;
XX KW receptor for HA mediated mobility; immunosuppressive; cytostatic.
XX KW conjugate; rheumatoid arthritis; scleroderma; liver fibrosis; cancer;
XX OS Rattus norvegicus.
XX PN WO200180899-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-CA00533.
XX PR 20-APR-2000; 2000US-198613P.
XX PA (CANG-) CANGENE CORP.
XX PI Woloski BMR, Williams AM, Sereda TJ, Wiebe DJ;
XX DR WPI; 2002-075094/10.
XX DR N-PSDB; AAS17498.
XX XX
XX PT Protein conjugates that selectively target certain tissues and organs
XX PT useful for treating and preventing various diseases, comprises
XX PT glucose-aminoglycan-targeting domain conjugated to a therapeutic
XX PT protein -
XX XX
XX PS Claim 6; Page 116-117; 121pp; English.
XX CC The invention relates to a conjugate comprising an hyaluronic acid (HA)
XX CC -binding protein e.g. RHAMM (receptor of HA mediated mobility) or peptide
XX CC contiguous with, or coupled to a polypeptide conjugated to a therapeutic
XX CC agent, and the polynucleotides encoding them. Also included is a method
XX CC for preparation of the HA-binding protein by inserting a first nucleotide
XX CC sequence encoding a HA-binding protein directly linked to a second
XX CC nucleotide sequence encoding a therapeutic protein into a suitable
XX CC vector, expressing the vector in an acceptable host, purifying conjugate
XX CC molecule from host or expression medium. The composition is useful for
XX CC altering in vivo the distribution of a therapeutic agent comprising
XX CC administering the composition to the animal where conjugate molecule will
XX CC distribute primarily in tissues and organs containing high levels of
XX CC endogenous HA and for treating mammal with a disorder where a diseased
XX CC tissue of the mammal contains high level of HA e.g. rheumatoid
XX CC arthritis, scleroderma, liver fibrosis and cancer. Lower therapeutic
XX CC dosages required also translates into lower immunogenicity of the
XX CC conjugated protein as compared to the native protein. As a result,
XX CC conjugates improve patient compliance and reduce direct and indirect
XX CC costs associated with the drug substance and its administration.
XX CC Conjugates allows for the use, where appropriate, of lower, safer,
XX CC dosages as compared to the conventional dosage requirements for the
XX CC unconjugated corresponding therapeutic agent. Conjugate molecules have an
XX CC increased half-life and potency, resulting in prolonged circulation of
XX CC the molecule, efficient distribution into the target tissues, and
XX CC increased bioavailability. The present sequence represents a RHAMM
XX CC protein.
XX SQ Sequence 713 AA;
Query Match 74.7%; Score 893; DB 23; Length 713;
Best Local Similarity 74.0%; Pred. No. 2e-56;
```

Matches 179; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

QY 1 QEKYDSWVQSLEDVTAQPEYSYKALTASIEDLKLENSLOEKAAGKAGNAEDVQHOILAT 60
 DB 442 QEKYSDTAQTLTDVTAQLESYKSSLTKEIEDLKLENLTOEKVAEKREVEDVQOQILTA 501
 QY 61 ESSNOEYVRMLDLQTKGALKETEIKETVTSFLOKITDLOQLKQOEEDFRKQLEDEGR 120
 DB 502 ESTNQEVAKVQDLQNSSTLKEAEIKETTSYLEKITDLOQLROQNEEDFRKQLEEGAK 561
 QY 121 KAEKENTTAELTEENKRWLLYBELYNTKPFQIOLDAFEVEKQALLNEHGAQOLNKI 180
 DB 562 MTEKETAVTELTMEINKWLLYELFDKTPFQQQLDAFEAEKQALLNEHGATQOLSKI 621
 QY 181 RDSYAKLIGHQNLKQIKHVVKLKDENSOLKSEVSKLRQOLAKKQSETKLQEEELNVLG 240
 DB 622 RDSYAQLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRQOLAKKQONELRLQGLDKALG 681
 QY 241 IK 242
 DB 682 IR 683

RESULT 14
 AAR43563
 ID AAR43563 standard; Protein; 476 AA.

XX
 AC AAR43563;
 XX
 DT 25-MAR-2003 (updated)
 DT 09-JAN-2003 (updated)
 DT 05-APR-1994 (first entry)
 XX
 DE Hyaluronan receptor.

XX
 KW Hyaluronan binding protein; HA; RHAMM; mediated motility; wound;
 KW healing; diagnosis; treatment; cell locomotion; tumour invasion;
 KW birth defects; inflammatory disorder; Alzheimer's disease; dementia;
 KW Parkinson's diseases; Huntington's disease; AIDS; diabetes; auto;
 KW immune diseases; corneal dysplasia; hypertrophy; surgery; burns;
 KW strokes; multiple sclerosis; depression; schizophrenia; CNJ;
 KW contraception; in vitro fertilisation; embryo development.

XX
 OS Rattus sp.
 XX
 PN WO9321312-Al.
 XX
 PD 28-OCT-1993.
 XX
 PF 13-APR-1993; 93WO-CA00158.
 XX
 PR 09-APR-1992; 92GP-0007949.
 XX
 PA (MANI-) MANITOBA CANCER TREATMENT & RES FOUND.
 PA (UYMA-) UNIV MANITOBA.
 XX
 XX Turley EA;
 PI
 DR WPI; 1993-351722/44.
 DR N-PSDB; AAQ51212.
 XX

PT DNA encoding hyaluronan receptor - used to produce proteins and
 PT antibodies for alteration of cell locomotion
 XX
 PS Claim 7; Fig 23; 89pp; English.
 XX

CC The sequence is that encoded by a cDNA clone encoding the hyaluronan
 CC receptor (HARC). The sequence was obtd. by screening a 3T3 library in
 CC lambda gt11 with antibodies to HARC. A clone of 1.9 kb was obtained
 CC and used to rescreen the library to obtain the full length, 2.9 kb
 CC clone. HA is down regulated in stationary normal cells and is only
 CC expressed in situations where cell motility is desired, e.g. in
 CC wound healing, in response to growth factors and in chemotaxis by

CC white blood cells. HA may be used for diagnosis and treatment of
 CC diseases involving cell locomotion, e.g. tumour invasion, birth
 CC defects, acute and chronic inflammatory disorders, Alzheimer's and
 CC other forms of dementia, AIDS, diabetes, autoimmune diseases, corneal
 CC dysplasias and hypertrophies, burns, surgical incisions and adhesions,
 CC strokes, multiple sclerosis, depression/schizophrenia related to
 CC neuronal growth and pain states involving nerve sprouting; also in CNJ
 CC and spinal cord regeneration, contraception, in vitro fertilisation and
 CC embryo development.
 CC See also AAR46548-51.
 CC (Updated on 09-JAN-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 476 AA;

Query Match 74.6%; Score 891.5; DB 14; Length 476;
 Best Local Similarity 75.2%; Pred. No. 1.6e-56;
 Matches 182; Conservative 20; Mismatches 39; Indels 1; Gaps 1;

QY 1 QEKYDSWVQSLEDVTAQPEYSYKALTASIEDLKLENSLOEKAAGKAGNAEDVQHOILAT 60
 DB 209 QEKYNDTAQSLRDVSAQLESYKSSLTKEIEDLKLENLTOEKVAAEKREVEDVQOQILTA 268
 QY 61 ESSNOEYVRMLDLQTKGALKETEIKETVTSFLOKITDLOQLKQOEEDFRKQLEDEGR 120
 DB 269 ESTNQEVAKVQDLQNSSTLKEAEIKETTSYLEKITDLOQLROQNEEDFRKQLEEGAK 328
 QY 121 KAEKENTTAELTEENKRWLLYBELYNTKPFQIOLDAFEVEKQALLNEHGAQOLNKI 180
 DB 329 TAERENVMTLTMEINKWLLY-ELYEKTQFQQQLDAFEAEKQALLNEHGATQOLNKI 387
 QY 181 RDSYAKLIGHQNLKQIKHVVKLKDENSOLKSEVSKLRQOLAKKQSETKLQEEELNVLG 240
 DB 388 RDSYAQLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRQOLAKKQONELRLQGLDKALG 447
 QY 241 IK 242
 DB 448 IR 449

RESULT 15
 AAB91998
 ID AAB91998 standard; Peptide; 42 AA.
 XX
 AC AAB91998;
 XX
 DT 22-JUN-2001 (first entry)
 XX
 DE Fibronectin fragment and fibrin related peptide SEQ ID NO:1174.
 XX
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidy; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX WO200069900-A2.
 PN
 PD 23-NOV-2000.
 XX
 PF 17-MAY-2000; 2000WO-US13576.
 XX
 PR 17-MAY-1999; 99US-0134406.
 PR 10-SEP-1999; 99US-0153406.
 PR 15-OCT-1999; 99US-0159783.
 XX
 XX (CONJ-) CONJUCHEM INC.
 PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 PI WPI; 2001-112059/12.
 XX
 DR WPI; 2001-112059/12.
 XX

PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT

XX Disclosure; Page 578; 733pp; English.

CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimide and maleimide groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.

XX Sequence 42 AA;

Query Match 17.6%; Score 210; DB 22; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 YAKLLGHONLKQKTHVVKLKDENSQKSEVSKLRCLAKKK 225
DB 1 YAKLLGHONLKQKTHVVKLKDENSQKSEVSKLRCLAKKK 42

RESULT 16

AA95451
ID AAB95451 standard; Protein; 436 AA.

XX AAB95451;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:17910.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

PS Claim 8; SEQ ID 17910; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 436 AA;

Query Match 17.5%; Score 209; DB 22; Length 436;
Best Local Similarity 24.8%; Pred. No. 3.3e-07;
Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;

QY 1 QEKYDSNVQSLSDY-----TAQFESYKA-----LTASEIEDLKLNSLQEKA 43
DB 166 KNEYNFKMROLEHVNDASABDPSKTPPHFQHLAKLLETQGEIEDGRASKTSLEHLV 225

QY 44 AKAGKNAEDVQHOILATESNQYVVRMLDLQ---TKSALKETEIKEITVSFLQKITDLQ 100
DB 226 TKLNEDREVKNAILRMKEQLREWNLRLESQQLIEKNWLLQGLDDIK--RQKENSQ 282

QY 101 N-----QLKQOEEDFRKQ-----LEDEGRKAENKNTTAELTEENKW 138
DB 283 NHPDNQKLKNEQESIKERLAKSKIVBEMLKADLEEVQSALYNKEMECILMTDEVERT 342

QY 139 RLIIYELYNKTKPFQILDAFEVEKQALINEHGAQEQOLANKIRDSYAKLGHONLKQKIK 198
DB 343 OTLESKAFQSKQLRSKLEEMYEERTISOEMMLRKQVECLAEENKLVGHQHLQKIQ 402

QY 199 HVVKLKDENSQKSEVSKLRQ---LAKKQSET 229

DB 403 YVVRLLKENVRLAEETEKLRANVFLKEKRSSES 436

RESULT 17

AAU79590
ID AAU79590 standard; Protein; 1388 AA.

XX AAU79590;

XX 24-SEP-2002 (first entry)

XX Human kinesin motor protein, Hskif15.

XX Human; enzyme; cytostatic; neuroprotective; kinesin motor protein;

XX Hskif15; microtubule stimulated ATPase activity; Hskif15M1A367;

XX Hskif15M1P401; kinesin; mitotic spindle; microtubule motor; antigen;

XX therapeutic; diagnosis; cancer; neurological disorder;

XX vesicular transport; atherosclerosis; tumour; abnormal wound healing;

XX inflammatory disorder; immune disorder; rheumatoid arthritis;

XX ocular angiogenic disease; glaucoma; cardiovascular disease;

XX hypertension; diastolic dysfunction; fungal disease; aspergillosis.

XX Homo sapiens.

PN US6391613-B1.
 XX 21-MAY-2002.
 XX 27-NOV-2000; 2000US-0723219.
 XX 04-JUN-1999; 99US-137423P.
 XX 17-MAY-2000; 2000US-0572191.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C, Sakowicz R, Wood KW;
 XX WPI; 2002-498776/53.
 XX N-PSDB; ABK86978.
 XX
 XX New nucleic acid encoding a human kinesin motor protein designated
 PT Hskif15 which has microtubule stimulated ATPase activity, for
 PT diagnosing and treating cancer, neurological disorders and disorders of
 PT vesicular transport -
 XX
 PS Claim 1; Fig 2; 21pp; English.
 XX
 CC The invention discloses an isolated nucleic acid encoding a human kinesin
 CC motor protein, Hskif15, which has microtubule stimulated ATPase activity,
 CC and two truncated versions, Hskif15MIA367 and Hskif15MIP401, retaining
 CC the amino terminus. Hskif15, a member of the kinesin superfamily, has
 CC been found to be essential for mitotic spindle formation. The predicted
 CC structure comprises an amino-terminal kinesin-like microtubule "motor"
 CC domain. The polynucleotides and polypeptides may be used therapeutically
 CC to diagnose and prevent or treat cancer, neurological disorders and
 CC disorders of vesicular transport. Examples of the conditions that can be
 CC treated include atherosclerosis, tumours, abnormal wound healing,
 CC inflammatory and immune disorders (such as rheumatoid arthritis), ocular
 CC angiogenic disease (such as glaucoma), cardiovascular disease (such as
 CC hypertension), diastolic dysfunction and fungal disease (such as
 CC aspergillosis). The polynucleotides and polypeptides may also be used to
 CC screen for modulators of Hskif15 and raise antibodies. The sequence
 CC presented is the human kinesin motor protein, Hskif15.
 XX
 SQ Sequence 1388 AA;
 Query Match 17.5%; Score 209; DB 23; Length 1388;
 Best Local Similarity 24.8%; Pred. No. 1.3e-06;
 Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;
 QY 1 QEKYDSMVQSLEDV-----TAQFESYKA-----LTASEIEDLKLENSLSQEA 43
 DB 1118 KNEYNFKMQLEHVMDSAEDPQSPKTPPHFQTHLAKLLETQOEIEDGRASKTSLEHLV 1177
 QY 44 AKAGKVAEDVQHOILATESNQEYVRMLDLQ---TKSALKETEIKEITVSPLOKITDQ 100
 DB 1178 TKLNEDREYKVAEILRMKEQLEMEINRLSQQLEKNMLLOGQLDDIK---ROKENSQ 1234
 QY 101 N-----OLKQOEEDFRKQ-----LEDEEGKAEKENTTAELTEINKW 138
 DB 1235 NHPDNQOLKNEQESIKERLAKSKIIVEMLKMKADLEEVQSALYNKMECLMTDVEVT 1294
 QY 139 RLIIYEELNNTKPFQIQDFAFEVQKALLNEHGAQEQQLNKIRDSYAKLIGHQNLKQKIK 198
 DB 1295 QTLESKAFQEKQLRSKLEEMYEERTSQEMEMLRKQVCELAENGKLVGHQNLHQKIQ 1354
 QY 199 HVVYKLKDNSOLSKSEVSKLRQCO---LAKKQOSET 229
 DB 1355 YVYRLAKENVRLAETEKLAEVFLKAEKRSSES 1388
 RESULT 18
 AAEE1400
 ID AAEE14400 standard; Protein; 1388 AA.
 XX
 AC AAEE14400;
 XX

DT 26-MAR-2002 (first entry)
 XX Human kinesin superfamily microtubule motor protein, Hskif15.
 DE Human; kinesin superfamily; microtubule motor protein; Hskif15; ATPase;
 XX adenosine triphosphatase; mitotic spindle; cancer; neurological disorder;
 KW vesicular transport; medicinal; veterinary; agricultural.
 KW Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Domain 32..391
 FT /label= Motor_domain
 XX
 XX WO200188118-A1.
 XX 22-NOV-2001.
 XX 17-MAY-2001; 2001WO-US16300.
 XX 17-MAY-2000; 2000US-0572191.
 XX (CYTO-) CYTOKINETICS INC.
 XX Beraud C, Sakowicz R, Wood KW;
 XX WPI; 2002-089854/12.
 XX N-PSDB; AAD23970.
 XX
 XX New isolated kinesin superfamily microtubule motor protein, Hskif15,
 XX useful for the diagnosis, treatment, or prevention of cancer,
 XX neurological disorders, and disorders of vesicular transport -
 XX
 PS Claim 1; Fig 2; 70pp; English.
 CC The present sequence is human kinesin superfamily microtubule motor
 CC protein, Hskif15. The motor protein shows microtubule stimulated
 CC adenosine triphosphatase (ATPase) activity and microtubule binding
 CC activity. The protein is essential for mitotic spindle formation.
 CC The Hskif15 polypeptide and polynucleotide are useful for the diagnosis,
 CC treatment or prevention of cancer, neurological disorders and disorders
 CC of vesicular transport. The polypeptide and polynucleotide are also
 CC useful in assays for identifying Hskif15 modulators, and in medicinal,
 CC veterinary, agricultural and research based applications.
 XX
 SQ Sequence 1388 AA;
 Query Match 17.5%; Score 209; DB 23; Length 1388;
 Best Local Similarity 24.8%; Pred. No. 1.3e-06;
 Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;
 QY 1 QEKYDSMVQSLEDV-----TAQFESYKA-----LTASEIEDLKLENSLSQEA 43
 DB 1118 KNEYNFKMQLEHVMDSAEDPQSPKTPPHFQTHLAKLLETQOEIEDGRASKTSLEHLV 1177
 QY 44 AKAGKVAEDVQHOILATESNQEYVRMLDLQ---TKSALKETEIKEITVSPLOKITDQ 100
 DB 1178 TKLNEDREYKVAEILRMKEQLEMEINRLSQQLEKNMLLOGQLDDIK---ROKENSQ 1234
 QY 101 N-----OLKQOEEDFRKQ-----LEDEEGKAEKENTTAELTEINKW 138
 DB 1235 NHPDNQOLKNEQESIKERLAKSKIIVEMLKMKADLEEVQSALYNKMECLMTDVEVT 1294
 QY 139 RLIIYEELNNTKPFQIQDFAFEVQKALLNEHGAQEQQLNKIRDSYAKLIGHQNLKQKIK 198
 DB 1295 QTLESKAFQEKQLRSKLEEMYEERTSQEMEMLRKQVCELAENGKLVGHQNLHQKIQ 1354
 QY 199 HVVYKLKDNSOLSKSEVSKLRQCO---LAKKQOSET 229
 DB 1355 YVYRLAKENVRLAETEKLAEVFLKAEKRSSES 1388
 RESULT 19

```
ABR48222
ID ABR48222 standard; Protein; 1388 AA.
XX
AC ABR48222;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated protein sequence SEQ ID NO:164.
XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US21338.
XX
PR 03-JUL-2001; 2001US-302814P.
XX
PR 03-AUG-2001; 2001US-310099P.
XX
PR 08-NOV-2001; 2001US-343705P.
XX
PR 13-NOV-2001; 2001US-350666P.
XX
PR 12-APR-2002; 2002US-372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
XX
DR N-PSDB; ACC51037.
XX
XX
PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with
PT a bladder cancer-associated polynucleotide or antibody
XX
XX
PS Claim 10; Page 287; 307pp; English.
XX
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications.
XX
XX
SQ Sequence 1388 AA;
Query Match 17.5%; Score 209; DB 24; Length 1388;
Best Local Similarity 24.8%; Pred. No. 1.3e-06;
Matches 68; Conservative 57; Mismatches 101; Indels 48; Gaps 7;
QY 1 QKYSVMQSLSDV-----TAQFSYKA---LTASEIEDLKLNSLSQKA 43
DB 1118 KNEYFKMQLEHVMDSAAEDPQSPKTPPHFQTHLAKLETQOEIEDGRKSTSLHLV 1177
QY 44 AXAGKNAEDVQHQLATESNQYVRMLDLQ---TKSALKETEIKETVTSFLOKITDQ 100
DB 1178 TKLNEDREVKNAEILRMKQLEMLNRLSQQSLIEKNWLLQQLDDIK--RQKNSDQ 1234
QY 101 N-----QLKQOEEDFRKQ-----LEDEGRKAEKENTTAELTETINKW 138
DB 1235 NHPDNOQLKNEBESIKERLAKSVIEMEMLKWKADLEEVQSALYNKEMECLRMTDEVERT 1294
QY 139 RLLYBELYNKTKPFQQLDAFEVQKALLNEHGAQQLNKIRDSYAKLLGHQNLKQKTK 198

ABR77430
ID ABR77430 standard; Protein; 795 AA.
XX
AC ABR77430;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human tumour marker protein se2-1.
XX
KW Human; tumour; cytostatic; cutaneous T cell lymphoma; CTCL; vaccine;
XX antigen-presenting cell; tumour-specific T cell.
XX
OS Homo sapiens.
XX
PN WO200238803-A2.
XX
PD 16-MAY-2002.
XX
PF 08-NOV-2001; 2001WO-DE04229.
XX
PR 08-NOV-2000; 2000DE-1055285.
XX
PA (DEKR-) DRUT KREBSFORSCHUNGSZENTRUM.
XX
PI Eichmueller S, Schadendorf D, Usener D;
XX
DR WPI; 2002-426959/45.
XX
DR N-PSDB; ABL58957.
XX
XX
PT Composition containing tumor-associated nucleic acid, useful for
PT diagnosis and treatment of tumors, especially cutaneous T cell
PT lymphoma, also derived proteins and antibodies -
XX
XX
PS Claim 5; Fig 7; 84pp; German.
XX
XX
CC The invention relates to a diagnostic composition containing at least one
CC of 23 nucleotide sequences (I, ABL58901-ABL58950) with altered expression
CC associated with tumors. (I), including antisense sequences and
CC ribozymes, also proteins (II, ABR77424-ABR77445) encoded by them and
CC antibodies specific for (II), are useful for diagnosis, monitoring and
CC treatment of tumors, especially cutaneous T cell lymphoma (CTCL).
CC (II) and antibodies to (II) are useful for vaccination. (II) can also be
CC used to prepare pre-loaded antigen-presenting cells or tumour-specific T
CC cells.
XX
XX
SQ Sequence 795 AA;
Query Match 16.0%; Score 191; DB 23; Length 795;
Best Local Similarity 23.3%; Pred. No. 1.4e-05;
Matches 79; Conservative 53; Mismatches 103; Indels 104; Gaps 12;
QY 1 QEKYDSMVQSLSDVTAQFSYKALTAS---BIEIDLK-----LENSSLQEKAAKAG 47
DB 402 EDQLKILTWELQKKSSELEEMTKLTNNKEVELEELKKVLGSEKTLTYENKQF-EKIAEEL 460
QY 48 KNAE-----DVQHQLATESNQYVRMLDLQTK---SALKETETI----- 85
DB 461 KGTEQELIGLLQAREKEVHDLQITAITTSEQYYSKVDKLTLENEKLNKNTLTSHC 520
QY 86 -----KEITVSFLOKITDQLQNLKQOEEDFRKLEDEGRKAEKENTTAELTETINKW 138
DB 521 NKLSLENKELT-----QETSDMTLELNQOEEDINNKKQEBRLMKQIELQETETQLRNEL 576
QY 139 RLLYBELYNKTKPFQQLDAFE-----VEKQALLNEHGAQQL 177
```

```

Db      : ||| | : || | : || | : || | : || | : || |
577 BYVREELKQKRDVCKLDKSENCNLRKQVENKNKYIEELQENKALKKGGTAESKQL 636
Qy      : ||| | : || | : || | : || | : || | : || |
178 N-----KIRDSVA-----KLLGHQNLKQKIKHVVKLKDENSOLK 211
Db      : ||| | : || | : || | : || | : || | : || |
637 NYVEIKVKNKLELESASAKQFGEITDTYQKEIEDKKISEENLLEVEKAKVIADAVKLQ 696
Qy      : ||| | : || | : || | : || | : || | : || |
212 SEVSKLRQ-----LAKKQSETKLOBELNKVLGI 241
Db      : ||| | : || | : || | : || | : || | : || |
697 KEIDK-RCQHKIAEMVALMEKHQYDKIIEERDSELGL 734

RESULT 21
AAG66581
ID AAG66581 standard; Protein; 976 AA.
XX
AC AAG66581;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human SCP-1 muten.
XX
KW Human; SCP-1; synaptonemal complex protein 1; cytostatic; gene therapy;
KW muten; mutant; cell transformation marker; diagnosis; cancer.
XX
OS Homo sapiens.
XX
PN US6232460-B1.
XX
PD 15-MAY-2001.
XX
PF 25-JUN-1998; 98US-0104324.
XX
PR 15-JUL-1997; 97US-0892702.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Tueraci O, Sahin U, Pfreundschuh M;
XX
DR WPI; 2001-342776/36.
XX
PT Novel nucleic acid molecule encoding a muten of SCP-1, useful as a
PT marker for cell transformation, and for identifying substances which
PT are immunoreactive and indicative of pathological conditions -
XX
PS Example 2; Column 15-20; 11pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule which encodes
CC a muten of synaptonemal complex protein 1 (SCP-1). The muten comprises
CC a sequence of 976 amino acids fully defined in the specification (with
CC the proviso that His at position 225 is replaced by Phe, and Gly at
CC position 226 is replaced by Gln). The SCP-1 muten is useful as a marker
CC for cell transformation, for diagnosis and treatment of cancer and for
CC screening substances which are immunoreactive and indicative of
CC pathological conditions. The present sequence is the muten of
CC SCP-1 provided in the specification.
XX
SQ Sequence 976 AA;
Query Match 16.0%; Score 191; DB 22; Length 976;
Best Local Similarity 23.3%; Pred. No. 1.8e-05;
Matches 79; Conservative 53; Mismatches 103; Indels 104; Gaps 12;
Qy 1 QEKYDSMVQSLSDVTAQFESYKALTAS---EIEDLK-----LENSLSQEAKAAG 47
Db 402 EDQLKILTMELQKSSLEEWTKTNKVELEELKXVLGKESKTLLEYNQF-EKIAEEL 460
Qy 48 KNAE-----DVQHQILATESNQYEVRLMLDLQTK---SALKETEI----- 85
Db 461 KGTEQELIGLQAREKEVHDLQLTAITTSQYKSEVKDLKTELENEKLNKTELTSHC 520
Qy 86 -----KEITVSPLOKITDLQNLQKQEBDFRQKLEDEGRKAEKENTTAELTEENKW 138

```

```

Db      : ||| | : || | : || | : || | : || | : || |
521 NKLSLENKELT-----QETSDMTLELKQKQEDINNKNKQBERMLKQIENLQETETLRNEL 576
Qy      : ||| | : || | : || | : || | : || | : || |
139 RLLVEELYNKTKPFIQIDAF-----VEKQALLNEHGAQEQOL 177
Db      : ||| | : || | : || | : || | : || | : || |
577 BYVREELKQKRDVCKLDKSENCNLRKQVENKNKYIEELQENKALKKGGTAESKQL 636
Qy      : ||| | : || | : || | : || | : || | : || |
178 N-----KIRDSVA-----KLLGHQNLKQKIKHVVKLKDENSOLK 211
Db      : ||| | : || | : || | : || | : || | : || |
637 NYVEIKVKNKLELESASAKQFGEITDTYQKEIEDKKISEENLLEVEKAKVIADAVKLQ 696
Qy      : ||| | : || | : || | : || | : || | : || |
212 SEVSKLRQ-----LAKKQSETKLOBELNKVLGI 241
Db      : ||| | : || | : || | : || | : || | : || |
697 KEIDK-RCQHKIAEMVALMEKHQYDKIIEERDSELGL 734

RESULT 22
ABP74709
ID ABP74709 standard; Protein; 976 AA.
XX
AC ABP74709;
XX
DT 03-FEB-2003 (first entry)
XX
DE Human SCP-1 protein SEQ ID NO:596.
XX
KW Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
KW T cell.
XX
OS Homo sapiens.
XX
PN WO200281646-A2.
XX
PD 17-OCT-2002.
XX
PF 04-APR-2002; 2002WO-US11101.
XX
PR 06-APR-2001; 2001US-282211P.
XX
PR 07-NOV-2001; 2001US-337017P.
XX
PR 07-MAR-2002; 2002US-363210P.
XX
PA (CTL1-) CTL IMMUNOTHERAPIES CORP.
XX
PI Simard JuL, Diamond DC, Liu L, Xie Z;
XX
DR WPI; 2003-067518/06.
XX
DR N-PSDB; ABQ83857.
XX
PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid
PT encoding the peptides, that are useful epitopes of target-associated
PT antigens -
XX
PS Claim 1; Page 180; 352pp; English.
XX
CC The present invention describes an isolated epitope (I) and an epitope
CC cluster. Also described is a vaccine or immunotherapeutic composition
CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for
CC treating an animal, by administering to an animal the vaccine or
CC immunotherapeutic composition. VC is also useful for evaluating, by
CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
CC ABP74173 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 976 AA;
Query Match 16.0%; Score 191; DB 24; Length 976;
Best Local Similarity 23.3%; Pred. No. 1.8e-05;

```

Matches 79; Conservative 53; Mismatches 103; Indels 104; Gaps 12;
QY 1 QEKYDSMVQSLDVTQAQFESYKALTAS---EIEDLK-----LENSLOFKAAG 47
Db 402 EDQKILTMELQKSSLEEMTKLTNNKEVELEELKKVLGKGTLLYENKQF-EKIAEEL 460
QY 48 KNAB-----DVQHILATESNQYVVMMLDIQTK---SALKETEI----- 85
Db 461 KGTEQELIGLQAREKEVHDLQITAITTSBOYYSKEVXDLKTELENEKLNKTELTSHC 520
QY 86 -----KEITVSLQKITDLOLQKQOEEDFRQKQDEDEGRKAEKENTTAELTEEINKW 138
Db 521 NKLSLENKELT-----QETSDMTLEKNQOEDINNKKQOEEMLKQIENLOETETOLNEL 576
QY 139 RLAYEELYNKTPQIOLDAPE-----VEKQALLNEHGAQEQOL 177
Db 577 EYVREELQKRDVEVKCKLDKXSEENCNLRKQVENKNKYIEBLOQENKALKKGTAEKQOL 636
QY 178 N-----KIRDSYA-----KLGHQNLKQKIKHVVKLKDENSOLK 211
Db 637 NVBEIKVKNLELESQKQFGEITDTYQKIEBKISEENLLEBEKAKVIADEAVKLQ 696
QY 212 SEVSKLRQ-----LAKKQSETKLOBELNKVLGI 241
Db 697 KEIDK-RCQHKIAEWALMEKHQYDKIIEERDSELGL 734

RESULT 23
AAW02258
ID AAW02258 standard; Protein; 1411 AA.
XX AC AAW02258;
XX DT 09-MAR-1997 (first entry)
XX DE Nucleolar/endosomal auto-antigen p162.
XX KW Auto-antibody; p162; rheumatic disease; antigen; diagnosis;
XX KW gene therapy.
XX OS Homo sapiens.
XX PN DE19515514-C1.
XX PD 12-SEP-1996.
XX PF 27-APR-1995; 95DE-1015514.
XX PR 27-APR-1995; 95DE-1015514.
XX PA (PRIV-) PRIVATES INST IMMUNOLOGIE & MOLEKULARGEN.
XX PI Renz M, Seelig HP;
XX DR WPI; 1996-403153/41.
XX DR N-PSDB; AAT58751.
XX PT DNA encoding nucleolar-endosomal auto-antigen - useful for exact
XX PT diagnosis of rheumatic disease, in gene therapy and for removal of
XX PT specific auto-antibodies
XX PS Claim 1; Fig 2; 15pp; German.
XX CC Transformed cells can be cultured to produce the antigen p162, for use
XX CC in exact (differential) diagnosis of rheumatic disease, i.e. they
XX CC can detect, in immunoassays, western blots, etc., rheumatism-
XX CC specific auto-antibodies. The antigen can be used therapeutically,
XX CC in the removal of auto-antibodies from the circulation, or when
XX CC coupled to a cytotoxin, the elimination of auto-antibody-
XX CC producing lymphocytes.
XX SQ Sequence 1411 AA;

Query Match 16.0%; Score 191; DB 17; Length 1411;
Best Local Similarity 27.7%; Pred. No. 2.7e-05;
Matches 74; Conservative 50; Mismatches 69; Indels 74; Gaps 12;
QY 1 QEKYDSMVQSLDVTQAQFESYKALTASIEIEDLKLENSLOFKAAGNAEDVQH 56
Db 715 KEKYSLEQKTELEBQIK-----KLEADSLEVKASKE-QALQDLQOORQLN 760
QY 57 -----ILATSSNQ-----EYVRL-IDLQTSKALKEITEIVSFLOKITDLOLQKQ 106
Db 761 TDLELATSLKQLEMEKEIVSVSTRDLQKKS-----EALSTKQKLT-----KOE 806
QY 107 EEDFRKQLEDEGRKAEKENTTAELTEEINKWELLYEELYNKTPQIOLDAPEVEKQAL 166
Db 807 EE---KQILKQDPETLSQET-----KIQHEELNNRIQTVTTELQKVNKEAL 851
QY 167 LNEHGAQEQOLNKIRDSY-----AKLGHQNLKQKIKHVVKLKDENSOLK 211
Db 852 MTELSTVKDKLSKVSDSLKNSKSEFEKNGKAAILDLEKTKELKHQLOVQVOMENT-LK 910
QY 212 SEVSKLRQCLAKKQKQSETKLOBELNKV 238
Db 911 -EQELKKSLEKEKEASHQKLEINSM 936
RESULT 24
ABB61144
ID ABB61144 standard; Protein; 1690 AA.
XX AC ABB61144;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 10224.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li FWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL05247.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 10224; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB16176-AB13051), expressed DNA
XX CC sequences (AB101840-AB16175) and the encoded proteins
XX CC (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

```

XX
SQ Sequence 1690 AA;
Query Match 15.8%; Score 189; DB 22; Length 1690;
Best Local Similarity 27.1%; Pred. No. 4.7e-05;
Matches 75; Conservative 54; Mismatches 106; Indels 42; Gaps 11;

QY 2 EKYSMVQSLEDVTAQFESYKALTASEIEDLKLENSLSQEKAA---KAGKNAEDVQHQL 58
Db 629 EQIRELNQQLDEVITQLNVQADSSALDMLRIQKEGTEERKSTLLEKTEKELVQSQA 688
QY 59 ATPESSNQYVRMLDLQTKSALKETEIKETIVSFLOKI-----TDLQNLKQEF-EDFR 111
Db 689 KTLNDKQLEQKQISDLK-QLAEQKLVREMTENAINQIQLEKSETEQQLALKQNELEDFQ 747
QY 112 KQLEDEGR---KAERKENTTAELTEERINKRWLLVEELYNKT---KPFQIOLDAFEVEKQ 164
Db 748 KKQSESEVHHQEIQAQNTQKDFELVESGESLKKQQQLQKTLGHEKIQAALEELKKEKE 807
QY 165 ALLNEHGAABQL-NKIRDSYAKL----LGHNLKQKI-----KHVVKLKDNSQLK 211
Db 808 TIIEKQELQOLQSKSAESSEALKVQVQLEQLQQQAASGESEKTVAKLHDSISQLK 867
QY 212 SEVSKLRQQL-----AKKKQSET---KLQELNK 237
Db 868 SQAETQSELKSTQNLLEAKSKQLAANGSLSEEA 904

RESULT 25
ABB61173
ID ABB61173 standard; Protein; 1690 AA.
XX
AC ABB61173;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10311.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
FI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-58DB; ABL05276.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 10311; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

```

```

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 1690 AA;
Query Match 15.8%; Score 189; DB 22; Length 1690;
Best Local Similarity 27.1%; Pred. No. 4.7e-05;
Matches 75; Conservative 54; Mismatches 106; Indels 42; Gaps 11;
QY 2 EKYSVMVQSLIEDVTAQFESYKALTASEIEDKLKLENSLSQKAA--KAGKNAEDVQHQIL 58
DB 629 EQIRELNQQQDEVTQTQNVOKADSSALDDMLRQKGETEKSTLLEKTELVSQKQAA 688
QY 59 ATBESSNQEYVRMLDLQTKSALKETEIKETIVSFLOKI-----TDLQNLKQOE-EDPR 111
DB 689 KTLNDKLEQLEKQISDLK-QLAEQEKLVRMTENAINQIQLEKESIEQQALKQNELEDFQ 747
QY 112 KQLEDBEGR---KAENKENTATLITEINKWRLLIYBELYNKT---KPFQOLDAFEVKQ 164
DB 748 KQOSESEVHLQETKQAQNTQKDFELVESGESLKKLQQQLKQLEKTLGHEKLAALBELKKEKE 807
QY 165 ALANEHGAAQEQ-LNKIRDSYAKI---LGHQNLKQKI-----KHVVVKLKDENSOLK 211
DB 808 TTIKEKEQELQLOKSAESALKVVQVQLEQLOQQAASGEGSKTVAKLHDEISOLK 867
QY 212 SEYSKLRQCOI-----AKKQOSET---KLQBELNK 237
DB 868 SQAEETQSELKSTQSNLEAKSKQLEAANGSLEEBACK 904
RESULT 26
ID ABG21233
AB ABG21233 standard; Protein; 1948 AA.
AC ABG21233;
XX
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #21224.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS85420.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX Claim 20; SEQ ID No 51592; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The

```


OS Lycopersicon esculentum.
 XX W0200061615-A2.
 PN
 XX
 XX
 PD 19-OCT-2000.
 KW
 XX
 XX 12-APR-2000; 2000WO-US09723.
 XX
 XX 12-APR-1999; 99US-0128900.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Harder PA, Meier I;
 PI
 XX WPI; 2000-679464/66.
 DR N-PSDB; AAA95812.
 XX
 XX Nucleic acid fragments from tobacco, corn, soybean and rice, encoding
 PT proteins that are homologs to the MAR binding filament-like protein 1
 PT (MFPI), useful for development of novel phenotypes -
 XX
 XX Claim 5; Page 54-56; 62pp; English.
 XX
 XX The present sequence is LemFP1, the matrix attachment region (MAR)
 CC binding filament-like protein 1 (MFPI) from tomato. MFPI
 CC has features of a novel anchor protein that most likely connects
 CC chromatin via MAR DNA with the nuclear envelope and nuclear filament
 CC proteins. MFPI nucleic acids and proteins may be used to better
 CC understand the mechanisms underlying this process so that the attachment
 CC of transgenes to the nuclear matrix may be used routinely to improve gene
 CC expression. They may be used to study MFPI expression, leading to the
 CC creation of novel developmental phenotypes that may be beneficial for
 CC crop growth and development. In addition, if the reduction in expression
 CC of one of the genes leads to a growth or developmental defect in the
 CC plant, this gene can be used as a novel herbicide target.
 XX
 XX Sequence 717 AA;
 SQ
 Query Match 14.8%; Score 177; DB 21; Length 717;
 Best Local Similarity 25.3%; Pred. No. 0.00012;
 Matches 71; Conservative 60; Mismatches 94; Indels 56; Gaps 13;
 QY 1 QEKYDSWQSLDVTAFQESYKALTA--SEIEDLKLENSSLQEKAAKAGKNAEDVOH-- 55
 DB 428 QBSLENSRSESDITVQLQLDLCLEAEVSKLQME---LEETPASLQRINIDETKHS 484
 QY 56 QILATE-----SSNOEYVRM-----LIDLQTSALKETEIKE-- 87
 DB 485 ELAARLTTTLLKKTNEEMHTMSDELVAENRDSLQTELVDVYKAEHTANELKQEK 544
 QY 88 -ITVFLQKITLQNLKQOEDFRKQLEDEGRKAEKENTAEITTEINKVRL-LYEEL 145
 DB 545 SIVATLEELKFLSQI-TREKELRKSLEDE----LEK---ATESUDEINRNVLAAEEL 596
 QY 146 YNKTFPQQLDAFEVEKQALLNEHGAQOBLNKIRDSYAKLL---GHNLKQKIKHV 201
 DB 597 ELATNSLSLEDERVLRGSVEFKQISQEAQENLEDASLVNKLKERESELEKRAK--- 653
 QY 202 KLDNSQLKSEVKLRQLCAKQKQSETKLQBELNKVLGK 242
 DB 654 KLEDEMAAKGEILRLRSQINSVK---APVEDEKVVAGEK 691
 RESULT 29
 ID ABB78804
 XX ABB78804 standard; Protein; 336 AA.
 AC ABB78804;
 XX
 XX 29-JUL-2002 (first entry)
 DT
 XX Myosin tail Myosin protein SEQ ID NO:33.
 DE
 XX

Human; NOVX; cytostatic; antiarteriosclerotic; cardiovascular; lymphoma;
 anti-diabetic; immunosuppressive; neuroprotective; gene therapy; cancer;
 cardiomyopathy; atherosclerosis; cell signal processing; diabetes; AIDS;
 metabolic pathway modulation; neoplastic; neurological disorder; asthma;
 adenocarcinoma; prostate cancer; uterus cancer; immune response;
 Crohn's disease; multiple sclerosis; Graft versus host disease.
 Unidentified.
 OS
 XX W0200230974-A2.
 XX
 XX 18-APR-2002.
 XX
 XX 12-OCT-2001; 2001WO-US31922.
 XX
 XX 12-OCT-2000; 2000US-240113P.
 XX 16-OCT-2000; 2000US-240625P.
 XX 16-OCT-2000; 2000US-240637P.
 XX 16-OCT-2000; 2000US-240648P.
 XX 16-OCT-2000; 2000US-240662P.
 XX 16-OCT-2000; 2000US-240669P.
 XX 16-OCT-2000; 2000US-240703P.
 XX 16-OCT-2000; 2000US-240732P.
 XX 16-OCT-2000; 2000US-241190P.
 XX 18-JAN-2001; 2001US-262455P.
 XX (CURA-) CURAGEN CORP.
 XX (MILL/) MILLET I.
 PA
 PA Grosse WM, Alsobrook JP, Lepley DM, Burgess CE, Mishra V;
 PI Kekuda R, Li L, Padigaru M, Shimkets RA, Zernhusen BD, Spytek KA;
 PI Edinger S, Gerlach V, MacDougall J, Stone D, Gunther E;
 PI Ellerman K;
 XX
 XX WPI; 2002-444172/47.
 XX
 XX New NOVX polypeptides and polynucleotides, useful for treating or
 PT preventing a NOVX-associated disorder or a pathological state in a
 PT subject, particularly a human, e.g. cardiomyopathy, atherosclerosis,
 PT cancer or diabetes -
 XX
 XX Disclosure; Page 27; 227pp; English.
 PS
 XX The present invention describes novel human proteins designated NOVX
 CC (where X is 1, 2a, 2b, 2c, 2d, 3, 4, 5, 6a, 6b, 7, 8, or 9). NOV1 is a
 CC tyrosine-protein kinase 6-like protein; NOV2a-d are keratin 4-like
 CC proteins; NOV3 is a collagen-like protein; NOV4 is a cystatin B-like
 CC protein; NOV5 is a serotonin receptor-like protein; NOV6a and NOV6b are
 CC cold inducible glycoprotein 30-like proteins; NOV7 is a matrilin-2-like
 CC protein; NOV8 is a leukocyte surface antigen (CD53)-like protein; and
 CC NOV9 is a tyrosine kinase-like protein. NOVX sequences have cytostatic,
 CC antiarteriosclerotic, cardiovascular, anti-diabetic, immunosuppressive,
 CC and neuroprotective activities, and can be used in gene therapy. The
 CC NOVX sequences can be used in therapeutics, particularly for treating,
 CC preventing or alleviating a NOVX-associated disorder or a pathological
 CC state in a subject, particularly a human. These disorders include
 CC cardiomyopathy, atherosclerosis, a disorder related to cell signal
 CC processing and metabolic pathway modulation or diabetes. The NOVX
 CC sequences are also useful for determining the presence of or
 CC predisposition to a disease associated with altered levels of NOVX
 CC polypeptide or nucleic acid, particularly cancer. The NOVX sequences are
 CC especially useful in therapeutic or prophylactic applications for
 CC neoplastic or neurological disorders, and in the treatment of
 CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
 CC response, AIDS, asthma, Crohn's disease, multiple sclerosis or Graft
 CC versus host disease. The present sequence represents a myosin tail
 CC protein sequence which is given in a domain analysis comparison
 CC with the human NOV2 protein from the present invention.
 XX
 XX Sequence 336 AA;
 SQ
 Query Match 14.4%; Score 171.5; DB 23; Length 336;
 Best Local Similarity 22.7%; Pred. No. 0.00013;

	Matches	64;	Conservative	48;	Mismatches	93;	Indels	77;	Gaps	8;
Qy	1	QEQYDSVMVQSL	EDYTAQFESYKALTA	GBIEDKL	ENSLQEA	KAAGNAEDVQHQILAT	60			
Db	15	QVKLDELQRLNDLTSQ	-----	-----	-----	KSFQSENSDLTRQLEAEAAQV	53			
Qy	61	ESSNQEVYVRMLLD	QTSALKETIKETIVSF	LQKITD	LQNLQOEED	---PRKOLEDE	117			
Db	54	--SNLSKLKLSQ	LESQLEAKSLEES	-----	-----	RERANLQALQLEHDLDSLRQLEEE	105			
Qy	118	EGRKAEKENTTAEL	TEINKWRLIYE	-----	-----	ELYNKTKPPQIQ	155			
Db	106	SEAKAELERQLSK	ANBIQWRSPFESGAL	RABELELKKLNQKISELE	EEAAEANA	165				
Qy	156	LDAPFEVEKQAL	-----	-----	-----	LNEHGAQEQNLKIRDSYAKLGHQNLKOKI	197			
Db	166	CDSLEKTKSR	LQSELEDLQIELE	RANAAAASELEKQQRNFDKILA	---EWKRKVDELQAE	LD 223				
Qy	198	---KHVVVLK	CDENSLKSEVSKLRC	QALKKQSETKLQEE	LN 236					
Db	224	TAQEAARNLS	TELPRLKNLELDELQ	VEALRRENKLNLODE	IEH 265					

RESULT 30

RESOL 30
ABB57354
ID ABB57354 standard; Protein; 1388 AA.

AA
DT 07-MAR-2002 (first entry)

XX DE Mouse ischaemic condition related protein sequence SEQ ID NO:993.

XX	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW	vasospastic ischaemia; ischaemic condition; ischaemic disease.

AA
OS
Mus musculus.

AA PN WO200188188-A2.

XX
PD
22-NOV-2001.

18-MAY-2001; 2001WO-JP04192.

XX
PR 18-MAY-2000; 2000JP-0145977.

XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

WPI; 2002-034733/04.

DR N-PSDB; ABI99859.

Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes -

XX PS Claim 2: Page 2517-2523; 2690pp; English.

The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI999202 to ABI99912, encoding the protein sequences in AB57020 to AB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR primers for a mouse ischaemic condition related sequence,

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 06:09:28 ; Search time 517 Seconds
(without alignments)
87.056 Million cell updates/sec

Title: US-09-978-309A-74

Perfect score: 1195
Sequence: 1 QEKYDSMVQSLDVTQAQFES.....KKKQSETKLOEELNKVLGIK 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1195	100.0	242	11	US-09-978-309A-74
2	1193	99.8	407	11	US-09-978-309A-78
3	1193	99.8	709	12	US-10-256-250-15
4	1193	99.8	725	11	US-09-978-309A-47
5	1193	99.8	725	12	US-10-256-250-14
6	1101.5	92.2	352	11	US-09-978-309A-83
7	1090	91.2	221	11	US-09-978-309A-76
8	922.5	77.2	676	12	US-10-258-250-16
9	917	76.7	631	11	US-09-978-309A-48
10	913	76.4	333	11	US-09-978-309A-73
11	913	76.4	476	11	US-09-978-309A-77
12	902.5	75.5	476	11	US-09-978-309A-79
13	869	72.7	435	11	US-09-978-309A-80
14	842	70.5	221	11	US-09-978-309A-75
15	209	17.5	1388	15	US-10-146-473-82

Sequence 596, App
Sequence 99, Appl
Sequence 33, Appl
Sequence 7646, Ap
Sequence 102, App
Sequence 2035, Ap
Sequence 5, Appl
Sequence 7, Appl
Sequence 4, Appl
Sequence 12, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 6, Appl
Sequence 5, Appl
Sequence 4, Appl
Sequence 148, App
Sequence 81, Appl
Sequence 14, Appl
Sequence 3290, Ap
Sequence 2, Appl
Sequence 8290, Ap
Sequence 3, Appl
Sequence 1548, Ap
Sequence 43, Appl
Sequence 7611, Ap
Sequence 62, Appl
Sequence 711, App
Sequence 44, Appl
Sequence 305, App
Sequence 168, App
Sequence 5251, Ap
Sequence 12141, A
Sequence 419, App
Sequence 3240, App

16 191 16.0 976 12 US-10-117-937-596
17 178.5 14.9 1940 12 US-09-738-630-99
18 171.5 14.4 336 12 US-09-976-782-33
19 169 14.1 1881 12 US-10-032-585-7646
20 168.5 14.1 892 12 US-10-205-219-102
21 168.5 14.1 975 12 US-10-094-749-2035
22 163 13.6 1379 12 US-10-205-219-5
23 160.5 13.4 1286 14 US-10-017-216-7
24 160.5 13.4 1958 14 US-10-028-948-4
25 160.5 13.4 2053 12 US-10-325-430-12
26 160.5 13.4 2053 14 US-10-017-216-2
27 160.5 13.4 2054 14 US-10-028-946-2
28 159.5 13.3 1441 12 US-10-412-897-3
29 159.5 13.3 1597 14 US-10-017-216-6
30 159.5 13.3 1641 14 US-10-017-216-5
31 159.5 13.3 2055 14 US-10-017-216-4
32 159 13.3 1312 12 US-10-393-602-148
33 157 13.1 32 11 US-09-978-309A-81
34 155 13.0 434 10 US-09-866-582-14
35 153 12.8 2099 15 US-10-128-714-3290
36 153 12.8 2285 10 US-09-932-183A-2
37 153 12.8 2405 15 US-10-128-714-8290
38 152 12.7 590 12 US-10-211-060-3
39 151.5 12.7 1742 12 US-10-012-697-1548
40 151 12.6 1203 15 US-10-097-340-43
41 150.5 12.6 1948 12 US-10-032-585-7611
42 150.5 12.6 1960 12 US-10-236-031B-62
43 149 12.5 374 9 US-09-925-302-711
44 148.5 12.4 621 12 US-10-316-253-44
45 148 12.4 689 14 US-10-108-605-305
46 147 12.3 677 9 US-09-745-763-168
47 147 12.3 996 9 US-09-815-242-5251
48 147 12.3 1009 9 US-09-815-242-12141
49 146.5 12.3 1979 15 US-10-205-823-419
50 146 12.2 1001 15 US-10-128-714-3240

ALIGNMENTS

RESULT 1

US-09-978-309A-74
; Sequence 74, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrik, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to Injury and Other Proliferating Cell Disorders Regulated by Hyaladerin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/585,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapien

Query Match 100.0%; Score 1195; DB 11; Length 242;
Best Local Similarity 100.0%; Pred. No. 2.2e-79;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEKYDSMVQSLDVTQAQFESYKALTASIEDLKLENSLQEKAAKAGNAEDVQHILAT 60

Db 1 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHQILAT 60
QY 61 ESSNOEYVRMLDLQTKSALKETEITVSFLQKITDLOLQKQOEEDFRKQLEDEGR 120
Db 61 ESSNOEYVRMLDLQTKSALKETEITVSFLQKITDLOLQKQOEEDFRKQLEDEGR 120
QY 121 KAEKENTTAELTEINKWRLLYBELYNKTPFOIQDADFVEVEKQALLNEHGAQOQLNKI 180
Db 121 KAEKENTTAELTEINKWRLLYBELYNKTPFOIQDADFVEVEKQALLNEHGAQOQLNKI 180
QY 181 RDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRCLAKKKQSETKLQELNKKVLG 240
Db 181 RDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRCLAKKKQSETKLQELNKKVLG 240
QY 241 IK 242
Db 241 IK 242

RESULT 2
US-09-978-309A-78
; Sequence 78, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978, 309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-78

Query Match 99.8%; Score 1193; DB 11; Length 407;
Best Local Similarity 99.8%; Pred. No. 5.6e-79;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHQILAT 60
Db 125 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHQILAT 184
QY 61 ESSNOEYVRMLDLQTKSALKETEITVSFLQKITDLOLQKQOEEDFRKQLEDEGR 120
Db 185 ESSNOEYVRMLDLQTKSALKETEITVSFLQKITDLOLQKQOEEDFRKQLEDEGR 244
QY 121 KAEKENTTAELTEINKWRLLYBELYNKTPFOIQDADFVEVEKQALLNEHGAQOQLNKI 180
Db 245 KAEKENTTAELTEINKWRLLYBELYNKTPFOIQDADFVEVEKQALLNEHGAQOQLNKI 304
QY 181 RDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRCLAKKKQSETKLQELNKKVLG 240
Db 305 RDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRCLAKKKQSETKLQELNKKVLG 364
QY 241 IK 242
Db 365 IK 366

RESULT 3
US-10-256-250-15
; Sequence 15, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/MXA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-15

Query Match 99.8%; Score 1193; DB 12; Length 709;
Best Local Similarity 99.6%; Pred. No. 1.1e-78;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHQILAT 60
Db 427 QEKYDSMVOSLEDVTAQFESYKALTASEIEDLKLENSLQEKAAKAGKNAEDVQHQILAT 486
QY 61 ESSNOEYVRMLDLQTKSALKETEITVSFLQKITDLOLQKQOEEDFRKQLEDEGR 120
Db 487 ESSNOEYVRMLDLQTKSALKETEITVSFLQKITDLOLQKQOEEDFRKQLEDEGR 546
QY 121 KAEKENTTAELTEINKWRLLYBELYNKTPFOIQDADFVEVEKQALLNEHGAQOQLNKI 180
Db 547 KAEKENTTAELTEINKWRLLYBELYNKTPFOIQDADFVEVEKQALLNEHGAQOQLNKI 606
QY 181 RDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRCLAKKKQSETKLQELNKKVLG 240
Db 607 RDSYAKLLGHQNLKQIKHVVKLKDENSOLKSEVSKLRCLAKKKQSETKLQELNKKVLG 666
QY 241 IK 242
Db 667 IK 668

RESULT 4
US-09-978-309A-47
; Sequence 47, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978, 309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-47

Query Match 99.8%; Score 1193; DB 11; Length 725;
Best Local Similarity 99.6%; Pred. No. 1.1e-78;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEKYDSMVQSLDVTQAQFESYKALTASIEIDLKLENSLSQEKAAKAGKNAEDVQHQILAT 60
DB 443 QEKYDSMVQSLDVTQAQFESYKALTASIEIDLKLENSLSQEKAAKAGKNAEDVQHQILAT 502

QY 61 ESSNOEYVRMLDLQTSKALKETEITVSFLQKITDLOQLKQOEEDFRKQLEDEGR 120
DB 503 ESSNOEYVRMLDLQTSKALKETEITVSFLQKITDLOQLKQOEEDFRKQLEDEGR 562

QY 121 KAEKENTTAELTEINKWRLLYEELYNKTKPQOLDFAFEVKQALLNEHGAQEQOLNKI 180
DB 563 KAEKENTTAELTEINKWRLLYEELYNKTKPQOLDFAFEVKQALLNEHGAQEQOLNKI 622

QY 181 RDSYAKLLGHQNLKQIKHVVKLKDENSQKSEVSKLRQOLAKKKQSETKIQBELNKVLG 240
DB 623 RDSYAKLLGHQNLKQIKHVVKLKDENSQKSEVSKLRQOLAKKKQSETKIQBELNKVLG 682

QY 241 IK 242
DB 683 IK 684

RESULT 5
US-10-256-250-14
; Sequence 14, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/MXA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 14
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-14

Query Match 99.8%; Score 1193; DB 12; Length 725;
Best Local Similarity 99.6%; Pred. No. 1.1e-78;
Matches 241; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QEKYDSMVQSLDVTQAQFESYKALTASIEIDLKLENSLSQEKAAKAGKNAEDVQHQILAT 60
DB 443 QEKYDSMVQSLDVTQAQFESYKALTASIEIDLKLENSLSQEKAAKAGKNAEDVQHQILAT 502

QY 61 ESSNOEYVRMLDLQTSKALKETEITVSFLQKITDLOQLKQOEEDFRKQLEDEGR 120
DB 503 ESSNOEYVRMLDLQTSKALKETEITVSFLQKITDLOQLKQOEEDFRKQLEDEGR 562

QY 121 KAEKENTTAELTEINKWRLLYEELYNKTKPQOLDFAFEVKQALLNEHGAQEQOLNKI 180
DB 563 KAEKENTTAELTEINKWRLLYEELYNKTKPQOLDFAFEVKQALLNEHGAQEQOLNKI 622

QY 181 RDSYAKLLGHQNLKQIKHVVKLKDENSQKSEVSKLRQOLAKKKQSETKIQBELNKVLG 240
DB 623 RDSYAKLLGHQNLKQIKHVVKLKDENSQKSEVSKLRQOLAKKKQSETKIQBELNKVLG 682

QY 241 IK 242
DB 683 IK 684

RESULT 6
US-09-978-309A-83
; Sequence 83, Application US/09978309A

Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladerin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-83

Query Match 92.2%; Score 1101.5; DB 11; Length 352;
Best Local Similarity 99.1%; Pred. No. 2.1e-72;
Matches 224; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 QEKYDSMVQSLDVTQAQFESYKALTASIEIDLKLENSLSQEKAAKAGKNAEDVQHQILAT 59
DB 125 QEKYDSMVQSLDVTQAQFESYKALTASIEIDLKLENSLSQEKAAKAGKNAEDVQHQILAT 184

QY 60 TESSNOEYVRMLDLQTSKALKETEITVSFLQKITDLOQLKQOEEDFRKQLEDEGR 119
DB 185 TESSNOEYVRMLDLQTSKALKETEITVSFLQKITDLOQLKQOEEDFRKQLEDEGR 244

QY 120 KAEKENTTAELTEINKWRLLYEELYNKTKPQOLDFAFEVKQALLNEHGAQEQOLNKI 179
DB 245 KAEKENTTAELTEINKWRLLYEELYNKTKPQOLDFAFEVKQALLNEHGAQEQOLNKI 304

QY 180 IRDSYAKLLGHQNLKQIKHVVKLKDENSQKSEVSKLRQOLAKKK 225
DB 305 IRDSYAKLLGHQNLKQIKHVVKLKDENSQKSEVSKLRQOLAKKK 350

RESULT 7
US-09-978-309A-76
; Sequence 76, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladerin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-76

Query Match 91.2%; Score 1090; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 8.2e-72;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KALTASEIEDLKLENSLOEKAAGKNAEDVQHOILATSSNQEVYRMLDLQTSALK 81
DB 1 KALTASEIEDLKLENSLOEKAAGKNAEDVQHOILATSSNQEVYRMLDLQTSALK 60

QY 82 ETEIKITVSFLOKITDNLQNLQKQBEDFRKQLEDEGRKAERNTTABELTEINKWRL 141
DB 61 ETEIKITVSFLOKITDNLQNLQKQBEDFRKQLEDEGRKAERNTTABELTEINKWRL 120

QY 142 YEELYNKTRFPQOLDFAFEVKQALLNEHGAQEOQLNKIRDSYAKLLGHONLKQKIKHV 201
DB 121 YEELYNKTRFPQOLDFAFEVKQALLNEHGAQEOQLNKIRDSYAKLLGHONLKQKIKHV 180

QY 202 KLDENSQKSEVSKLRCOLAKKQSETKLOEELNKVLGK 242
DB 181 KLDENSQKSEVSKLRCOLAKKQSETKLOEELNKVLGK 221

RESULT 8

US-10-256-250-16
; Sequence 16, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/WKA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 16
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-16

Query Match 77.2%; Score 922.5; DB 12; Length 676;
Best Local Similarity 79.3%; Pred. No. 4.3e-59;
Matches 192; Conservative 1; Mismatches 0; Indels 49; Gaps 1;

QY 1 QEKYDSMVQSLSDVTAQFESYKALTASEIEDLKLENSLOEKAAGKNAEDVQHOILAT 60
DB 443 QEKYDSMVQSLSDVTAQF----- 461

QY 61 ESSNQEVYRMLDLQTSALKETEIKITVSFLOKITDNLQNLQKQBEDFRKQLEDEGR 120
DB 462 -----RMLDLQTSALKETEIKITVSFLOKITDNLQNLQKQBEDFRKQLEDEGR 513

QY 121 KAEKNTTABELTEINKWRLLYEELYNKTRFPQOLDFAFEVKQALLNEHGAQEOQLNKI 180
DB 514 KAEKNTTABELTEINKWRLLYEELYNKTRFPQOLDFAFEVKQALLNEHGAQEOQLNKI 573

QY 181 RDSYAKLLGHONLKQKIKHVVKLDENSQKSEVSKLRCOLAKKQSETKLOEELNKVLG 240
DB 574 RDSYAKLLGHONLKQKIKHVVKLDENSQKSEVSKLRCOLAKKQSETKLOEELNKVLG 633

QY 241 IK 242
DB 634 IK 635

RESULT 9

US-09-978-309A-48
; Sequence 48, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony

; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-48

Query Match 76.7%; Score 917; DB 11; Length 631;
Best Local Similarity 76.4%; Pred. No. 1e-58;
Matches 185; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 QEKYDSMVQSLSDVTAQFESYKALTASEIEDLKLENSLOEKAAGKNAEDVQHOILAT 60
DB 363 QEKYNDTAQSLRDVTAQLESYKSTLKEIETDKLENLTQEKVMAAEKSVEDVQOQILTA 422

QY 61 ESSNQEVYRMLDLQTSALKETEIKITVSFLOKITDNLQNLQKQBEDFRKQLEDEGR 120
DB 423 ESTNQEVYARNVQDLQNRSTLKEEIEKITSSPLEKITDLKNQLRQODEDFRQLEBKGR 482

QY 121 KAEKNTTABELTEINKWRLLYEELYNKTRFPQOLDFAFEVKQALLNEHGAQEOQLNKI 180
DB 483 TAEKENVMTELTMWINKWRLLYEELYEYKTPFQOQLDAFEAKQALLNEHGAQEOQLNKI 542

QY 181 RDSYAKLLGHONLKQKIKHVVKLDENSQKSEVSKLRCOLAKKQSETKLOEELNKVLG 240
DB 543 RDSYAKLLGHONLKQKIKHVVKLDENSQKSEVSKLRCOLAKKQSETKLOEELNKVLG 602

QY 241 IK 242
DB 603 IR 604

RESULT 10

US-09-978-309A-73
; Sequence 73, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Mus musculus

US-09-978-309A-73

Query Match 76.4%; Score 913; DB 11; Length 333;
Best Local Similarity 76.0%; Pred. No. 9.3e-59;
Matches 184; Conservative 20; Mismatches 38; Indels

[illegible]

RESULT 11

```

US-09-978-309A-77
; Sequence 77, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-978-309A-77

```

Query Match	76.4%;	Score 913;	DB 11;	Length 476;
Best Local Similarity	76.0%;	Pred. NO. 1.4e-58;		
Matches 184;	Conservative 20;	Mismatches 38;	Indels 0;	Gaps 0
Qy	1	QEKYDSMVQSLSDVTAQPESSYKALTAASEIEDLKLNSLQEKRAAKAGKNAEDVQHQILAT	60	
Db	208	QEKYNDTAQSLRDVSAQLSEYSSTLKEITEDLKLNLTLQEKVAMAEKSVEDVQOQILTA	267	
Qy	61	ESSNCEYVRMLDLDTKALKETETKEITVTSFLQKTTDLQNLKQOEDDFRQKLEDEGR	120	
Db	268	ESTNQEYARMQDQLNRSITLKKEEIEKEITSSPLEKITDLKNQURQOEDDFRQKLEEKGR	327	
Qy	121	KAEKENTTAELTEENKWRLLLYEELYNKTKPQIQILDAFEVSKOALLNBEHGAAGQOLNK	180	
Db	328	TAEKENVMTLFWENKWRLLYEELYEYKTPPQQQIDAFEAKQOALLNBSHGATQEOLNKI	387	
Qy	181	RDSYAKLLGHQNLKQIKRKHVVKLKDENSOLKGEVSKRLQCLAKKKQSETKLOEELNKVLG	240	

Db 388 RDSYAQLLGHQNLKQIKHVVKLXDNSQLKXSEVSKLRSQLVKRRQNELRQGELOKALG 447

Qy 241 IK 242
|
448 IR 449

Db

RESULT 12

US-09-978-309A-79
; Sequence 79, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladerhin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/695,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-978-309A-79

Query Match	75.5%;	Score	902.5;	DB	11;	Length	476;
Best Local Similarity	76.0%;	Pred.	No. 8.2e-58;				
Matches	184;	Conservative	19;	Mismatches	38;	Indels	1;
Gaps	1;						
Qy	1	QKYSWMSQSLDVTQAQFESYKALTASIEDLKLNSLSQEAARAKAGNAEDVQHILAT	60				
Db	209	QKRYNTAQSRLDVTQAQLESYKSSPTKEIEDLKLNLTLQEKVAMAKSVEDVQQQILTA	268				
Qy	61	ESSNQEVYEMLDLQTSALKAKETHETKVTVSFLOKITTDLONOLQOQESDFRKOLEDESGR	120				
Db	269	ESTNQIYASWVDLQNRSTLKEEETKE-TSSFKEKITDLKNQLRQOQEDFRKQLEERKGR	327				
Qy	121	KAEKENTTAELTEENKRWLYEELYNKTKPFQIQDAFEVKQALLNEHGAQEQQLNKI	180				
Db	328	TAKEKNVWTELTWENKRWLYEELIYEKTKPFQOQOLDAPAEKQALLNEHGAQEQQLNKI	387				
Qy	181	RDSYAKLLGHONLKQIKRHHVKLKDENSQKLSYVKLRCLQAKKKQSQETKLQEBLNKVLG	240				
Db	388	RDSYAQLLGHONLKQIKRHHVKLKDENSQKLSYVKLRCLQAKKKQSQETKLQEBLNKVLG	447				
Qy	241	IK 242					
Db	448	IR 449					

RESIST. 13

RES001.13
 US-09-978-309A-80
 ; Sequence 80, Application US/09978309A
 ; Publication No. US20030100490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cruz, Tony
 ; APPLICANT: Pastrak, Aleksandra
 ; APPLICANT: Turley, Eva A.
 ; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Responses to
 ; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
 ; TITLE OF INVENTION: Hyaladherin and Hyaluronans
 ; FILE REFERENCE: 033352-010
 ; CURRENT APPLICATION NUMBER: US/09/978.309A


```
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-978-309A-80

Query Match      72.7%; Score 869; DB 11; Length 435;
Best Local Similarity 77.1%; Pred. No. 2e-55;
Matches 175; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

Qy 1 QEKYDSMVQSLDVTQAQESYKALTASHIEDKLENSLOEKAAGKNAEDVQHQIILAT 60
Db 209 QEKYNDTAQSLDVSQAQESYKSTLKEIEDKLENTLOEKVMAEKSVEDVQOQIILTA 268

Qy 61 ESSNOEYVRMLDLCTKSALKETEIKETVSPLOKITDLOLQKQOEDDFRQKLEDERGR 120
Db 269 ESTNOEYARMVQDLQNRSTKEEBKEITSSFLEKITDLOLQKQOEDDFRQKLEBKGR 328

Qy 121 KAEKENTTAETLTKKRWLLYEEYLYNKKTPQIQDFAFEVEKQALLNEHGAQOQLNKKI 180
Db 329 TAEKENVMTLTKKRWLLYEEYLYNKKTPQIQDFAFEVEKQALLNEHGAQOQLNKKI 388

Qy 181 RDSYAKLGHQNLKOKIKHVVKLKDNSOLKSEVSKRLQOLAKKQOS 227
Db 389 RDSYAQLLGHQNLKOKIKHVVKLKDNSOLKSEVSKRLQOLAKKQON 435

RESULT 14
US-09-978-309A-75
; Sequence 75, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-978-309A-75

Query Match      70.5%; Score 842; DB 11; Length 221;
Best Local Similarity 76.9%; Pred. No. 8.2e-54;
Matches 170; Conservative 18; Mismatches 33; Indels 0; Gaps 0;

Qy 22 KALTASEIEDKLENSLOEKAAGKNAEDVQHQIILATSSNQEVYVRMLDLQTSALK 81
Db 1 KSSTLKEIEDKLENTLOEKVMAEKSVEDVQOQIILTAESTNQEVYVRMLDLQNRSTLK 60
Qy 82 ETSIKEITVSPLOKITDLOLQKQOEDDFRQKLEDEBGRKAEKENTTAETLTKKRWLL 141
```

```
Db 61 ESEIKEITSSFLEKITDLOLQKQOEDDFRQKLEBKGRKAEKENVMTLTKKRWLL 120
Qy 142 YEELYNKKTPQIQDFAFEVEKQALLNEHGAQOQLNKKIRDSYAKLGHQNLKOKIKHVV 201
Db 121 YEELYXKTPQIQDFAFEVEKQALLNEHGAQOQLNKKIRDSYAKLGHQNLKOKIKHVV 180
Qy 202 KLKDNSOLKSEVSKRLQOLAKKQOSETKLOEELNKVLGK 242
Db 181 KLKDNSOLKSEVSKRLQOLAKKQOSETKLOEELNKVLGK 221

RESULT 15
US-10-146-473-82
; Sequence 82, Application US/10146473
; Publication No. US20030108889A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 82
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-82

Query Match      17.5%; Score 209; DB 15; Length 1388;
Best Local Similarity 24.8%; Pred. No. 6.2e-07;
Matches 68; Conservative 57; Mismatches 10; Indels 48; Gaps 7;

Qy 1 QEKYDSMVQSLDVTQAQESYKALTASHIEDKLENSLOEKA 43
Db 1118 KNEYNFKRQLEHVWDSAAEDPQSPKTPPHQTHLAKLETQEQEIEDGRASKTSLEHLV 1177
Qy 44 AKAGNAEDVQHQIILATSSNQEVYVRMLDLQ---TKSALKETEIKETVSPLOKITDLO 100
Db 1178 TKLNEDEVQNAEILRMKEQOLREMLNRLSQQLEKWLQGGDDIK---RQKENSQD 1234
Qy 101 N-----QLKQOEDDFRQK-----LEDEBGRKAEKENTTAETLTKKRWLL 138
Db 1235 NHPDQQLNQEESIEKRLAKSVEMLMKADLEEVQSLNLYNKECLEMTDEVERT 1294
Qy 139 RLLYEELYNKKTPQIQDFAFEVEKQALLNEHGAQOQLNKKIRDSYAKLGHQNLKOKIK 198
Db 1295 QTLKSAFOEKQELSKLEEMVEERERTSQEMEMLRKQVECLAENGKLVGHQNLFOKIQ 1354
Qy 199 HVVKKLKDENSOLKSEVSKRLQO---LAKKQOSET 229
Db 1355 YVVRLLKENVRLAEETEKLRANVFLKKEKRSSES 1388

RESULT 16
US-10-117-937-596
; Sequence 596, Application US/10117937
; Publication No. US2003022039A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPITOPE SEQUENCES
```



```
Db 54 --SNLSKLSQLESQLEAKSLEES-----RERANLQALQLEHDLDSLREQLEEE 105
QY 118 EGKAEKNTVAETEEINKRWLLYE-----ELYNKTKPFIQ 155
Db 106 SEAKELERQSKANAEIQWRSKFESGALRAEBELKKNQKISELEBAEAAANAK 165
QY 156 LDAFEVEKQAL-----LNEHGAAQOINKIRDSYAKLGHQNLKQKI 197
Db 166 CDSLEKTSRLQSELELDQIELERANAAASELEKQKNFKILA--EWRKRVDELQABLD 223
QY 198 ---XHVVKLKDENSOLKSEVSKLROLAKKQOSETKLOEELN 236
Db 224 TAQREARNLSTELFKNELBELKQOVALRRENKNLODEIH 265

RESULT 19
US-10-032-585-7646
; Sequence 7646, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7646
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Candida albicans
; NAME/KEY: MISC FEATURE
; LOCATION: (1881)..(1881)
; OTHER INFORMATION: X=any amino acid
US-10-032-585-7646

Query Match 14.1%; Score 169; DB 12; Length 1881;
Best Local Similarity 21.7%; Pred. No. 0.00071; Indels 96; Gaps 12;
Matches 70; Conservative 62; Mismatches 94;

QY 9 QSLEDVTAQFESYKALTASEIEDLKLENSLOFKAAGK-----NAEDVQHQI 57
Db 886 REFKNLTVEFNTKDYELQINNLSKNNEFKQKINELSKKTESLTDNKNFAKQLEKL 945
QY 58 LATESNQBYVRMLDLQT-----XSALKETE-----IKEITVSPFKITDILQNLKQOE 107
Db 946 RDTENNEHMDKLRSASVAYNDLKAKSESEETVKAKEELETLTSTKIDNLEKELBQO 1005
QY 108 -----EDFRKQLEDE--EGKAEKE--NTAEI----- 131
Db 1006 SKNHELGOLQNTDSTNEKF--KELEDELKSKSKNSKELSSQNSSELQKLEKTEKDLOAK 1064
QY 132 TBEINKWRL-----LYBELYNKTKPFIQOLDAPFEVQKALLNEHGAAQOINKIRDSY--- 184
Db 1065 DBEIDKKAETKSNIDNLNSEISSLSQSKLEABESHSSTKDBHSSLSNKLKKEBYENT 1124
QY 185 -----AKLLGHO-----NLQKIKHVVKLKDENSOLKS-----F 213
Db 1125 KTSWIAKLAKIEHKATDEITKTKHTDLOEEHAKQKQSPESERNDIKSNLDKANKE 1184
QY 214 VSKLROLAKKQOSETKLOEEL 235
Db 1185 LSDNREKLSNLEKEKTELNKL 1206
```

RESULT 20

US-10-205-219-102

; Sequence 102, Application US/10205219

```
; Publication No. US20030138803A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brookbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WI-A-018200
; CURRENT APPLICATION NUMBER: US/10/205,219
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Myosin heavy chain
US-10-205-219-102

Query Match 14.1%; Score 168.5; DB 12; Length 892;
Best Local Similarity 20.8%; Pred. No. 0.00032; Indels 121; Gaps 10;
Matches 74; Conservative 54; Mismatches 107;

QY 8 VQSLIEDVTAQFESYKALTASEIEDLKLENSLOKAKAGNAEDVQHQIILATESNOEY 67
Db 164 IDNLQVRQKLEKEKSEMKMEIDDLASNMVEISKGNLEKWCRTLEDQVSELKTKBEEQ 223
QY 68 VRMLDLQTSALKETETKEIT-----VSFLOKITDLOLQKQOED- 109
Db 224 QRLNELTAQGRGLQTESGEYSRQLEDKDSLVSQLSRGKQAFQOIEELKQLESEVAK 283
QY 110 -----FRKQLEDEBEGKAEKENTTAETBEINKWRLY-----BEL 145
Db 284 SALAHALOSSRHCDLLREQVEEHEQAKAELORAMSKANSEVAQWRTKYETDAIQTEEL 343
QY 146 YNKTKPFIQOLD-----APEVFKQALLNE-----HGAAQOINKIR 181
Db 344 BEAKKLAQRLQDAEBEHVAVNAKASLEKTKQRLQNEVEDLMIDVERTNAACAALDKQ 403
QY 182 DSYAKLIG-----HQNLKQ-----KIKHVVK-----LKDENSOLK 211
Db 404 RNFDKILAEWKQVEETHAELEASQKESRSLSLELFKIKNAVEESLDOLETIKRENKLIQ 463
QY 212 SEVSKLRCQLAK-----KKQ-----SETKLOEELNKVLGK 242
Db 464 QEISDLTEQIAEGGKRIHELEKIKKQIQEKSSELQALAEAEASLEHEEGKILRIQ 519

RESULT 21
US-10-094-749-2035
; Sequence 2035, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKI, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
```



```

QY      5 DSMVQSLEDVTAFESVKALTASE-----IEDKLENSSLOEKAAKAGNA 50
DB      797 DSKIRSLERIVSELSEANKLAANSSLTQRNMKAQEEMISELROOKFYLETQAGKLEAQN 856
QY      51 EDVOHQILATESSNQEVYRMLLDIOTKSALKETEIKETITVSFLOKITDLQNQLKQOEEDF 110
DB      857 RKEEQLEKISHODSHDKNRLLLETRLRVLSLEHEBQKLELKRQLTQLSLOERESQL 916
QY      111 -----RKOLEDE-EGRKAEXENTTAELTEINKWRLLYEELYNKTKPFQIOLDAFEVEK 163
DB      917 TALQAARAALLESOLRQAKTLEETTAAEBEIQALTAHRDEIQRK-----FDALRNSC 969
QY      164 QALLNEHGAQEOQLNKIRDSYAKULGHON--LKQIKHVVKLDKENSOLKSEVSKLRCOL 221
DB      970 TVITD---LEEQLNLQTEDNAE-LNNQNFYLSKQLEASGANDEIVQLRSEVDHLRREI 1024
QY      222 AKKKQSETKLQOEL 235
DB      1025 TEREMOLT SQKOTM 1038

RESULT 26
US-10-017-216-2
; Sequence 2, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Pr
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2

Query Match          13.4%; Score 160.5; DB 14; Length 2053;
Best local Similarity 23.6%; Pred. No. 0.0033;
Matches 60; Conservative 48; Mismatches 111; Indels 35; Gaps 7;

QY      5 DSMVQSLEDVTAFESVKALTASE-----IEDKLENSSLOEKAAKAGNA 50
DB      797 DSKIRSLERIVSELSEANKLAANSSLTQRNMKAQEEMISELROOKFYLETQAGKLEAQN 856
QY      51 EDVOHQILATESSNQEVYRMLLDIOTKSALKETEIKETITVSFLOKITDLQNQLKQOEEDF 110
DB      857 RKEEQLEKISHODSHDKNRLLLETRLRVLSLEHEBQKLELKRQLTQLSLOERESQL 916
QY      111 -----RKOLEDE-EGRKAEXENTTAELTEINKWRLLYEELYNKTKPFQIOLDAFEVEK 163
DB      917 TALQAARAALLESOLRQAKTLEETTAAEBEIQALTAHRDEIQRK-----FDALRNSC 969
QY      164 QALLNEHGAQEOQLNKIRDSYAKULGHON--LKQIKHVVKLDKENSOLKSEVSKLRCOL 221
DB      970 TVITD---LEEQLNLQTEDNAE-LNNQNFYLSKQLEASGANDEIVQLRSEVDHLRREI 1024
QY      222 AKKKQSETKLQOEL 235
DB      1025 TEREMOLT SQKOTM 1038

RESULT 27
US-10-028-946-2
; Sequence 2, Application US/10028946
; Publication No. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan

```

```
; APPLICANT: Miranda, Maricar
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US20020123622a1el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2054
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-2

Query Match      13.3%; Score 160.5; DB 14; Length 2054;
Best Local Similarity 23.6%; Pred. No. 0.0033;
Matches 60; Conservative 48; Mismatches 111; Indels 35; Gaps 7;

QY 5 DSMVQSLDVTAQFESYKALTASE-----IEDLKLSLSLOEKAAGKAGNA 50
Db 813 DSKIRSLQRIVELSEANKLAANSLFTQRMKAQEMISELRQKPFYLETQAGKLEAQN 872
QY 51 EDVQHQLATESNOQVYRMILLDTQTSALKETKEITVSFLQKITDQNLQKQBEDF 110
Db 873 RKLBEQLEKISHODSHQSKRLELETLRLREVSLEHEEQKLEKQLTELQLSQERESQL 932
QY 111 -----RKOLEDE-EGRKAEKENTTAELTEINKRWLLYEELYNKTKPFQIQLDAPEVK 163
Db 933 TALQARAALLESQRLQAKTELETTAEAEIEQALTAHRDEIQRK-----FDALRNSC 985
QY 164 QALLNEHGAQAQOLNKIRDSYAKLLGHQN--LKQIKHVVKLKDENSQKSEVSKLRQCL 221
Db 986 TVITD-----LEEQLNQLTDEANAE-LNNQNFYLSKQLDASGANDEIVQLRSEVDHLRREI 1040
QY 222 AKKQSETKQBEL 235
Db 1041 TEREMQLTSQKQTM 1054

RESULT 28
US-10-412-897-3
; Sequence 3, Application US/10412897
; Publication No. US20030220224A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES ENCODING THE HUMAN CITRON KINASE
; FILE REFERENCE: D0193 NP
; CURRENT APPLICATION NUMBER: US/10/412,897
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: U.S. 60/372,745
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-412-897-3

Query Match      13.3%; Score 159.5; DB 12; Length 1441;
Best Local Similarity 23.6%; Pred. No. 0.0025;
Matches 60; Conservative 48; Mismatches 111; Indels 35; Gaps 7;

QY 5 DSMVQSLDVTAQFESYKALTASE-----IEDLKLSLSLOEKAAGKAGNA 50
Db 198 DSKIRSLQRIVELSEANKLAANSLFTQRMKAQEMISELRQKPFYLETQAGKLEAQN 257
QY 51 EDVQHQLATESNOQVYRMILLDTQTSALKETKEITVSFLQKITDQNLQKQBEDF 110
Db 258 RKLBEQLEKISHODSHQSKRLELETLRLREVSLEHEEQKLEKQLTELQLSQERESQL 317
```

```
QY 111 -----RKOLEDE-EGRKAEKENTTAELTEINKRWLLYEELYNKTKPFQIQLDAPEVK 163
Db 318 TALQARAALLESQRLQAKTELETTAEAEIEQALTAHRDEIQRK-----FDALRNSC 370
QY 164 QALLNEHGAQAQOLNKIRDSYAKLLGHQN--LKQIKHVVKLKDENSQKSEVSKLRQCL 221
Db 371 TVITD-----LEEQLNQLTDEANAE-LNNQNFYLSKQLDASGANDEIVQLRSEVDHLRREI 425
QY 222 AKKQSETKQBEL 235
Db 426 TEREMQLTSQKQTM 439

RESULT 29
US-10-017-216-6
; Sequence 6, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPILLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prc
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-6

Query Match      13.3%; Score 159.5; DB 14; Length 1597;
Best Local Similarity 23.6%; Pred. No. 0.0029;
Matches 60; Conservative 48; Mismatches 111; Indels 35; Gaps 7;

QY 5 DSMVQSLDVTAQFESYKALTASE-----IEDLKLSLSLOEKAAGKAGNA 50
Db 354 DSKIRSLQRIVELSEANKLAANSLFTQRMKAQEMISELRQKPFYLETQAGKLEAQN 413
QY 51 EDVQHQLATESNOQVYRMILLDTQTSALKETKEITVSFLQKITDQNLQKQBEDF 110
Db 414 RKLBEQLEKISHODSHQSKRLELETLRLREVSLEHEEQKLEKQLTELQLSQERESQL 473
QY 111 -----RKOLEDE-EGRKAEKENTTAELTEINKRWLLYEELYNKTKPFQIQLDAPEVK 163
Db 474 TALQARAALLESQRLQAKTELETTAEAEIEQALTAHRDEIQRK-----FDALRNSC 526
QY 164 QALLNEHGAQAQOLNKIRDSYAKLLGHQN--LKQIKHVVKLKDENSQKSEVSKLRQCL 221
Db 527 TVITD-----LEEQLNQLTDEANAE-LNNQNFYLSKQLDASGANDEIVQLRSEVDHLRREI 581
QY 222 AKKQSETKQBEL 235
Db 582 TEREMQLTSQKQTM 595

RESULT 30
US-10-017-216-5
; Sequence 5, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPILLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prc
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
```

